

PATENT OFFICE  
JAPANESE GOVERNMENT



This is to certify that the annexed is a true copy of the following application as filed with this Office.

DATE OF APPLICATION : May 22, 2001

APPLICATION NUMBER : Patent Application 2001-152243

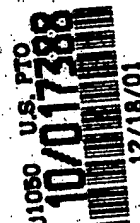
APPLICANT(S) : MONOLITH CO., LTD.

November 16, 2001

Commissioner, Patent Office

Kozo OIKAWA

日 本 国 特 許 庁  
JAPAN PATENT OFFICE



別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office

出 願 年 月 日

Date of Application:

2001年 5月22日

出 願 番 号

Application Number:

特願2001-152243

出 願 人  
Applicant(s):

株式会社モノリス

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2001年11月16日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造



出証番号 出証特2001-3101743

【書類名】 特許願

【整理番号】 MN-0130

【提出日】 平成13年 5月22日

【あて先】 特許庁長官殿

【国際特許分類】 H04B 14/00

【発明者】

    【住所又は居所】 東京都江戸川区西葛西 5 - 1 0 - 2 6

    【氏名】 品川 嘉久

【発明者】

    【住所又は居所】 千葉県千葉市稲毛区天台 4 - 2 - 1 2

    【氏名】 永嶋 浩樹

【特許出願人】

    【識別番号】 396001980

    【氏名又は名称】 株式会社モノリス

【先の出願に基づく優先権主張】

    【出願番号】 特願2000-395777

    【出願日】 平成12年12月26日

【代理人】

    【識別番号】 100105924

    【弁理士】

    【氏名又は名称】 森下 賢樹

    【電話番号】 0422-23-7415

【手数料の表示】

    【予納台帳番号】 054977

    【納付金額】 21,000円

【提出物件の目録】

    【物件名】 明細書 1

    【物件名】 図面 1

    【物件名】 要約書 1

特 2 0 0 1 - 1 5 2 2 4 3

【包括委任状番号】 9717729

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 画像補間方法と装置、および画像処理方法と装置

【特許請求の範囲】

【請求項 1】 2 枚のキーフレームからなる第 1 画像対と、それら 2 枚のキーフレーム間の第 1 対応点情報を取得する工程と、

2 枚のキーフレームからなる第 2 画像対と、それら 2 枚のキーフレーム間の第 2 対応点情報を取得する工程と、

第 1 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 1 軸と第 2 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 2 軸の位置関係、第 1 対応点情報、および第 2 対応点情報を利用して、中間フレームを補間によって生成する工程と、

を含むことを特徴とする画像補間方法。

【請求項 2】 前記第 1 軸と第 2 軸が同一直線上にこないよう第 1 画像対と第 2 画像対を定める請求項 1 に記載の方法。

【請求項 3】 第 1 画像対の 2 枚のキーフレームのうち 1 枚と、第 2 画像対の 2 枚のキーフレームのうち 1 枚を共通化し、第 1 軸と第 2 軸を二辺とする三角形を基礎に前記補間を行う請求項 1、2 のいずれかに記載の方法。

【請求項 4】 第 1 画像対の 2 枚のキーフレームと第 2 画像対の 2 枚のキーフレームを共通化せず、第 1 軸と第 2 軸を対向する二辺とする四辺形を基礎に前記補間を行う請求項 1、2 のいずれかに記載の方法。

【請求項 5】 前記中間フレームと第 1 画像対の 2 枚のキーフレームおよび第 2 画像対の 2 枚のキーフレームとの位置関係を取得する工程を含み、その位置関係をもとに前記補間を行う請求項 1 から 4 のいずれかに記載の方法。

【請求項 6】 第 1 画像対の 2 枚のキーフレームおよび第 2 画像対の 2 枚のキーフレームは、それぞれ同一視点から異なる時間において撮影された画像である請求項 1 から 5 のいずれかに記載の方法。

【請求項 7】 第 1 画像対の 2 枚のキーフレームおよび第 2 画像対の 2 枚のキーフレームは、それぞれ異なる視点から撮影された画像である請求項 1 から 5 のいずれかに記載の方法。

【請求項 8】 2 枚のキーフレームからなる第 1 画像対についてマッチング計算を行い、それら 2 枚のキーフレーム間の第 1 対応点情報を検出する工程と、

2 枚のキーフレームからなる第 2 画像対についてマッチング計算を行い、それら 2 枚のキーフレーム間の第 2 対応点情報を検出する工程と、

第 1 画像対の 2 枚のキーフレーム間において時間的または空間的に定まる第 1 軸と第 2 画像対の 2 枚のキーフレーム間において時間的または空間的に定まる第 2 軸の位置関係、第 1 対応点情報、および第 2 対応点情報を利用して、中間フレームを補間によって生成することを特徴とする画像補間方法。

【請求項 9】 前記検出する工程は、2 枚のキーフレームの間で画素単位のマッチング計算を行う請求項 8 に記載の方法。

【請求項 1 0】 前記検出する工程は、2 枚のキーフレームにおいてそれぞれ二次元的な探索を行って検出した特異点の対応をもとに画素単位のマッチング計算を行う請求項 9 に記載の方法。

【請求項 1 1】 前記検出する工程は、2 枚のキーフレームをそれぞれ前記特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得する請求項 1 0 に記載の方法。

【請求項 1 2】 複数のキーフレームを記憶するユニットと、

中間フレームの時間的または空間的な位置情報をキーフレームとの関係において取得するユニットと、

それぞれが 2 枚のキーフレームからなる第 1 画像対、第 2 画像対についてそれぞれ定まる対応点情報、および前記位置情報をもとに中間フレームを補間処理にて生成するユニットと、

を含み、第 1 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 1 軸と第 2 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 2 軸が同一直線上にこないようそれらの画像対が選定されることを特徴とする画像補間装置。

【請求項 1 3】 前記対応点情報を生成するマッチングプロセッサをさらに

含む請求項 1 2 に記載の装置。

【請求項 1 4】 第 1 画像対の 2 枚のキーフレームおよび第 2 画像対の 2 枚のキーフレームは、それぞれ同一視点から異なる時間において撮影された画像である請求項 1 2 に記載の装置。

【請求項 1 5】 第 1 画像対の 2 枚のキーフレームおよび第 2 画像対の 2 枚のキーフレームは、それぞれ異なる視点から撮影された画像である請求項 1 2 に記載の装置。

【請求項 1 6】 前記中間フレームの時間的または空間的な位置に関する指定を外部から入力するためのユーザインタフェイスをさらに含む請求項 1 2 から 1 5 のいずれかに記載の装置。

【請求項 1 7】 生成された中間フレームの時間的または空間的な位置に関する指定を外部から入力するためのユーザインタフェイスをさらに含む請求項 1 2 から 1 5 のいずれかに記載の装置。

【請求項 1 8】 2 枚のキーフレームからなる第 1 画像対と、それら 2 枚のキーフレーム間の第 1 対応点情報を取得する工程と、

2 枚のキーフレームからなる第 2 画像対と、それら 2 枚のキーフレーム間の第 2 対応点情報を取得する工程と、

第 1 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 1 軸と第 2 画像対の 2 枚のキーフレーム間に時間的または空間的に定まる第 2 軸の位置関係、第 1 対応点情報、および第 2 対応点情報を利用して、中間フレームを補間によって生成する工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【請求項 1 9】 キーフレーム間の対応点を記述した対応点ファイルを複数準備し、それらの間に混合処理を施すことにより新たな対応点ファイルを生成することを特徴とする画像処理方法。

【請求項 2 0】 生成された新たな対応点ファイルをもとに、前記キーフレーム間の中間フレームを補間によって生成する処理をさらに含む請求項 2 0 に記載の方法。

【請求項 2 1】 キーフレーム間の対応点を記述した対応点ファイルを複数

準備し、それらの間に混合処理を施すことにより新たな対応点ファイルを生成する処理をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、画像補間技術に関し、とくにデジタル画像補間方法と装置、および画像処理方法と装置に関する。

【0002】

【従来の技術】

インターネットの利用が広がるにつれ、当初は情報の提供が主たる用途であったネットワークビジネスが、オンラインショッピングへ急速に展開している。最近では、すでに強力な販売網を有するような大手企業がインターネットを通じた直販に次々と乗り出しており、何世紀にもわたって築きあげられてきた商取引の形態が大きな曲がり角にきているといっても過言ではない。

【0003】

こうした取引構造のリストラクチャリングともいうべき動きは、企業サイドからすれば中間コストの削減と、在庫レス、店舗レスによる効率経営の観点から歓迎され、一方、ユーザからは家庭に居ながらにして多数の商品情報にアクセスして望む商品を手に入れる簡便さが支持されている。PC（パーソナルコンピュータ）をはじめとする電子機器のさらなる普及と、商品提供者および商品需要者の双方に利する商形態である点に鑑みれば、オンラインショッピングが今後さらに急激な拡大を見せることに疑いはない。

【0004】

【発明が解決しようとする課題】

オンラインショッピングを体験してすぐに感ずるのは、商品をさらに手に取るように眺めたいという欲求である。実際に米国のある会社の調査では、商品を単に静止画像で見せる場合に比べ、回転させて擬似的に三次元で見せたとき、売上が数倍になるという結果が得られている。しかしながら、そのためにサイト側が



すべき準備は多大で、商品の回転表示する場合、ある程度滑らかに見せるためには、約 $1^{\circ}$ ごとに360枚の写真をとり、これをユーザの視点に応じて切り替える手法が知られている。

## 【0005】

サイトの問題はおくとしても、そうした多数の写真をユーザ端末に送信することは非常に厄介である。なぜなら、送信すべきデータ量はすぐに数メガから数十メガバイトになり、そのデータを毎回ダウンロードしなければならないとすれば、とてもユーザの気持ちに叶わない。

## 【0006】

本発明はこうした状況に鑑みてなされたものであり、その目的は、商品その他任意のオブジェクトを軽いデータで擬似的に三次元表示し、かつその際に高い画質を維持できる画像補間および処理技術の提供にある。

## 【0007】

## 【課題を解決するための手段】

本発明は画像補間または処理技術に関するが、その技術は必ずしも商品だけをその用途とするものではない。たとえば、映画等における画像補間、動画の圧縮効果もまったく同等に実現されるものであり、当然にそれも本発明の用途の範疇にある。

## 【0008】

本発明のある態様は、画像補間方法に関する。この方法は以下の工程を含む。

(1) 2枚のキーフレームからなる第1画像対と、それら2枚のキーフレーム間の第1対応点情報を取得する工程

(2) 2枚のキーフレームからなる第2画像対と、それら2枚のキーフレーム間の第2対応点情報を取得する工程

(3) 第1画像対の2枚のキーフレーム間に時間的または空間的に定まる第1軸と第2画像対の2枚のキーフレーム間に時間的または空間的に定まる第2軸の位置関係、第1対応点情報、および第2対応点情報を利用して、中間フレームを補間によって生成する工程。

## 【0 0 0 9】

(1) (2) において第1対応点情報、第2対応点情報は、ともにキーフレームどうしのマッチングによって求めてもよい。(3) において、第1軸と第2軸を利用して2軸補間 (bilinear interpolation) を行ってもよい。一例として、2つの視点  $p_1 (0, 0)$ 、 $p_2 (0, 100)$  から得たキーフレームを第1画像対とし、別の2つの視点  $p_3 (100, 0)$ 、 $p_4 (100, 100)$  から得たキーフレームを第2画像対とする。点  $p_1$  と  $p_2$  を結ぶ直線が第1軸、点  $p_3$  と  $p_4$  を結ぶ直線が第2軸に当たる。

## 【0 0 1 0】

中間フレームとして、視点  $p' = (50, 50)$  から見た画像を得る場合、まず第1画像対の間の対応点情報をもとに視点  $(0, 50)$  からのフレームを生成する。つぎに第2画像対の間の対応点情報をもとに視点  $(100, 50)$  からの別のフレームを生成する。つづいて、これらふたつのフレームに補間を施し、すなわちこの場合1:1で内分し、所望の中間フレームが生成される。ここで、縦横両方向の補間をするためには、一般には、第1軸と第2軸が同一直線上にこないよう第1画像対と第2画像対を定めればよい。

## 【0 0 1 1】

なお、この例では第1軸、第2軸は、それぞれ2枚のキーフレーム間に空間的に定まるが、時間的に定まる別の例もある。たとえば、ひとつの視点  $P$  から時刻  $t = t_0$  と  $t_1$  に得たふたつのキーフレームを第1画像対とし、別の視点  $Q$  から時刻  $t = t_0$  と  $t_1$  に得たふたつのキーフレームを第2画像対とする。この場合、第1画像対において  $(P, t_0)$  で定義される点と  $(P, t_1)$  で定義される点を結ぶ直線が第1軸となり、同様に第2画像対において  $(Q, t_0)$  で定義される点と  $(Q, t_1)$  で定義される点を結ぶ直線が第2軸となる。したがって、中間フレームを例えば点  $((P+Q)/2, (t_0+t_1)/2)$  からの画像とすれば、それぞれふたつの軸の中間的な画像を生成した後、これらを補間すればよい。以下、時間も空間も単に4次元のうちのパラメータとして特別に分けて考えることはしない。

## 【0 0 1 2】

第1画像対の2枚のキーフレームのうち1枚と、第2画像対の2枚のキーフレームのうち1枚を共通化し、第1軸と第2軸を二辺とする三角形を基礎に前記補間を行ってもよい。また、第1画像対の2枚のキーフレームと第2画像対の2枚のキーフレームを共通化せず、第1軸と第2軸を対向する二辺とする四辺形を基礎に前記補間を行ってもよい。

## 【0013】

中間フレームと第1画像対の2枚のキーフレームおよび第2画像対の2枚のキーフレームとの位置関係を取得する工程を含み、その位置関係をもとに前記補間を行ってもよい。この工程は、最初の例でいえば、中間フレームの視点位置を(50, 50)と定めることにあり、ここにユーザの意図が反映されてもよい。

## 【0014】

第1、第2対応点情報を検出する場合、その工程は、2枚のキーフレームにおいてそれぞれ二次元的な探索を行って検出した特異点の対応をもとに画素単位のマッチング計算を行ってもよい。

## 【0015】

その検出する工程は、2枚のキーフレームをそれぞれ前記特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得してもよい。

## 【0016】

本発明の別の態様は、画像補間装置に関する。この装置は、複数のキーフレームを記憶するユニットと、中間フレームの時間的または空間的な位置情報をキーフレームとの関係において取得するユニットと、それぞれが2枚のキーフレームからなる第1画像対、第2画像対についてそれぞれ定まる対応点情報、および前記位置情報をもとに中間フレームを補間処理にて生成するユニットとを含み、第1画像対の2枚のキーフレーム間に時間的または空間的に定まる第1軸と第2画像対の2枚のキーフレーム間に時間的または空間的に定まる第2軸が同一直線上にこないようそれらの画像対が選定される。この装置はさらに、対応点情報を生

成するマッチングプロセッサを含んでもよい。

【0017】

前述の特異点を用いるマッチング方法は、本出願人が先に特許第2927350号にて提案した技術（以下「前提技術」という）の応用であり、前記検出する工程に好適である。ただし、前提技術は縦横にわたる補間には触れていない。本発明の新しい技術により、例えば商品をいろいろな角度から眺めた画像を生成することができ、EC（電子商取引）などに好適である。

本発明の別の態様は、画像処理方法に関する。この方法は、キーフレーム間の対応点を記述した対応点ファイルを複数準備または取得し、それらの間に混合処理を施すことにより新たな対応点ファイルを生成するものである。「混合処理」は例えば2軸に関する補間（Bilinear interpolation）である。対応点ファイルの取得のために、後述のマッチングプロセッサを利用してもよい。こうして生成された新たな対応点ファイルをもとに、前記キーフレーム間の中間フレームを補間によって生成する処理をさらに含んでもよい。

【0018】

なお、本発明は前提技術を必須とはしない。また、以上の各構成、工程を任意に入れ替えたり、方法と装置の間で表現を一部または全部入れ替え、または追加したり、表現をコンピュータプログラム、記録媒体等に変更したものもまた、本発明として有効である。

【0019】

【発明の実施の形態】

はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第2927350号を得ている技術であり、本発明との組合せに最適である。本発明では、画像上にメッシュを設け、その格子点によって多数の画素を代表されるため、もともと前提技術のような画素単位のマッチング技術に対する適用効果が高いためである。ただし、実施の形態で採用可能な画像マッチング技術はこれに限られない。

【0020】

図 1 8 以降、前提技術を利用した画像データ符号化および復号技術を具体的に説明する。

〔前提技術の背景〕

ふたつの画像の自動的なマッチング、つまり画像領域や画素どうしの対応付けは、コンピュータビジョンやコンピュータグラフィックスにおける最も難しくかつ重要なテーマのひとつである。例えば、あるオブジェクトに関して異なる視点からの画像間でマッチングがとれれば、他の視点からの画像を生成することができる。右目画像と左目画像のマッチングが計算できれば、立体画像を用いた写真測量も可能である。顔の画像のモデルと他の顔の画像のマッチングがとれたとき、目、鼻、口といった特徴的な顔の部分を抽出することができる。例えば人の顔と猫の顔の画像間でマッチングが正確にとられたとき、それらの中割画像を自動的に生成することでモーフィングを完全自動化することができる。

【 0 0 2 1 】

しかし従来一般に、ふたつの画像間の対応点は人がいちいち指定しなければならず、多大な作業工数を要した。この問題を解消するために数多くの対応点自動検出方法が提案されている。例えば、エピポーラ直線を用いることによって対応点の候補の数を減らす考えがある。しかし、その場合でも処理はきわめて複雑である。複雑さを低減するために、左目画像の各点の座標は通常右目画像でもほぼ同じ位置にあると想定される。しかし、こうした制約を設けると、大域的特徴及び局所的特徴を同時に満たすマッチングをとることは非常に困難になる。

【 0 0 2 2 】

ボリュームレンダリングでは、ボクセルを構成するために一連の断面画像が用いられる。この場合、従来一般に、上方の断面画像における画素が下方の断面画像の同一箇所にある画素と対応すると仮定され、これらの画素のペアが内挿計算に用いられる。このようにきわめて単純な方法を用いるため、連続する断面間の距離が遠く、オブジェクトの断面形状が大きく変化する場合、ボリュームレンダリングで構築されたオブジェクトは不明瞭になりがちである。

【 0 0 2 3 】

立体写真測量法など、エッジの検出を利用するマッチングアルゴリズムも多い

。しかしこの場合、結果的に得られる対応点の数が少ないため、マッチングのとれた対応点間のギャップを埋めるべく、デイスパリティの値を内挿計算しなければならない。一般にあらゆるエッジ検出器は、それらが用いる局所的なウィンドウの中で画素の輝度が変化したとき、これが本当にエッジの存在を示唆するかどうかを判断することが難しい。エッジ検出器は、本来的にすべてハイパスフィルタであり、エッジと同時にノイズも拾ってしまう。

## 【 0 0 2 4 】

さらに別の手法として、オプティカルフローが知られている。二枚の画像が与えられたとき、オプティカルフローでは画像内のオブジェクト（剛体）の動きを検出する。その際、オブジェクトの各画素の輝度は変化しないと仮定する。オプティカルフローでは例えば  $(u, v)$  のベクトル場の滑らかさといった、いくつかの付加的な条件とともに、各画素の動きベクトル  $(u, v)$  を計算する。しかし、オプティカルフローでは画像間の大域的な対応関係を検出することはできない。画素の輝度の局所的な変化に注目するのみであり、画像の変位が大きい場合、システムの誤差は顕著になる。

## 【 0 0 2 5 】

画像の大域的な構造を認識するために、多重解像度フィルタも数多く提案されてきた。それらは線形フィルタと非線形フィルタに分類される。前者の例としてウェーブレットがあるが、線形フィルタは一般に、画像マッチングにはさして有用ではない。なぜなら、極値をとる画素の輝度に関する情報がそれらの位置情報とともに次第に不鮮明になるためである。図 1 (a) と図 1 (b) は顔の画像に対して平均化フィルタを適用した結果を示している。同図のごとく、極値をとる画素の輝度が平均化によって次第に薄れるとともに、位置も平均化の影響でシフトしていく。その結果、目（輝度の極小点）の輝度や位置の情報は、このような粗い解像度レベルで曖昧になり、この解像度では正しいマッチングを計算することができない。したがって、粗い解像度レベルを設けるのが大域的なマッチングのためでありながら、ここで得られたマッチングは画像の本当の特徴（目、つまり極小点）に正確に対応しない。より精細な解像度レベルで目が鮮明に現れたとしても、大域的なマッチングをとる際に混入した誤差は、もはや取り返しがつか

ない。入力画像にスムージング処理を加えることにより、テクスチャ領域のステレオ情報が落ちてしまうこともすでに指摘されている。

#### 【 0 0 2 6 】

一方、最近地形学の分野で利用されはじめた非線形フィルタとして一次元の「ふるい (sieve)」演算子がある。この演算子は、所定の大きさの一次元ウィンドウ内の極小値（または極大値）を選択することにより、縮尺と空間の因果関係を保存しながら画像にスムージング処理を加える。その結果得られる画像は元の画像と同じ大きさであるが、小さな波の成分が取り除かれているため、より単純になる。画像の情報を落とすという点で、この演算子は広い意味での「多重解像度フィルタ」に分類することはできるが、実際にはウェーブレットのように画像の解像度を変えながら画像を階層化するわけではなく（つまり狭い意味での多重解像度フィルタではなく）、画像間の対応の検出には利用できない。

#### 【 0 0 2 7 】

[前提技術が解決しようとする課題]

以上をまとめれば以下の課題が認められる。

1. 画像の特徴を正確に、かつ比較的簡単な処理で把握する画像処理方法が乏しかった。特に、特徴のある点に関する情報、例えば画素値や位置を維持しながら特徴を抽出できる画像処理方法に関する有効な提案が少なかった。
2. 画像の特徴をもとに対応点を自動検出する場合、一般に処理が複雑であるか、ノイズ耐性が低いなどの欠点があった。また、処理に際していろいろな制約を設ける必要があり、大域的特徴及び局所的特徴を同時に満たすマッチングをとることが困難だった。
3. 画像の大域的な構造または特徴を認識するために多重解像度フィルタを導入しても、そのフィルタが線形フィルタの場合、画素の輝度情報と位置情報が曖昧になった。その結果、対応点の把握が不正確になりやすかった。非線形フィルタである一次元ふるい演算子は画像を階層化しないため、画像間の対応点の検出には利用できなかった。
4. これらの結果、対応点を正しく把握しようとするれば、結局人手による指定に頼るほか有効な手だてがなかった。

## 【 0 0 2 8 】

前提技術はこれらの課題の解決を目的としてなされたものであり、画像処理の分野において、画像の特徴の的確な把握を可能にする技術を提供するものである。

## 【 0 0 2 9 】

[前提技術が課題を解決するための手段]

この目的のために前提技術のある態様は、新たな多重解像度の画像フィルタを提案する。この多重解像度フィルタは画像から特異点を抽出する。したがって、特異点フィルタともよばれる。特異点とは画像上特徴をもつ点をいう。例として、ある領域において画素値（画素値とは、色番号、輝度値など画像または画素に関する任意の数値を指す）が最大になる極大点、最小になる極小点、ある方向については最大だが別の方向については最小になるような鞍点がある。特異点は位相幾何学上の概念であってもよい。ただし、その他どのような特徴を有してもよい。いかなる性質の点を特異点と考えるかは、前提技術にとって本質問題ではない。

## 【 0 0 3 0 】

この態様では、多重解像度フィルタを用いた画像処理が行われる。まず検出工程において、第一の画像に対し、二次元的な探索を行って特異点が検出される。つぎに生成工程において、検出された特異点を抽出して第一の画像よりも解像度の低い第二の画像が生成される。第二の画像には第一の画像のもつ特異点が引き継がれる。第二の画像は第一の画像よりも解像度が低いため、画像の大域的な特徴の把握に好適である。

## 【 0 0 3 1 】

前提技術の別の態様は特異点フィルタを用いた画像マッチング方法に関する。この態様では、始点画像と終点画像間のマッチングがとられる。始点画像および終点画像とは、ふたつの画像の区別のために便宜的に与えた名称であり、本質的な違いはない。

## 【 0 0 3 2 】

この態様では、第一工程にて、始点画像に特異点フィルタを施して解像度の異



なる一連の始点階層画像が生成される。第二工程では、終点画像に特異点フィルタを施して解像度の異なる一連の終点階層画像が生成される。始点階層画像、終点階層画像とは、それぞれ始点画像、終点画像を階層化して得られる画像群をいい、それぞれ最低 2 枚の画像からなる。つぎに第三工程において、始点階層画像と終点階層画像のマッチングが解像度レベルの階層の中で計算される。この態様によれば、多重解像度フィルタによって特異点に関連する画像の特徴が抽出され、および／または明確化されるため、マッチングが容易になる。マッチングのための拘束条件は特に必要としない。

## 【 0 0 3 3 】

前提技術のさらに別の態様も始点画像と終点画像のマッチングに関する。この態様では、予め複数のマッチング評価項目のそれぞれに関して評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索する。総合評価式は、評価式の少なくともひとつに係数パラメータを掛けたうえでそれらの評価式の総和として定義してもよく、その場合、総合評価式またはいずれかの評価式がほぼ極値をとる状態を検出して前記パラメータを決定してもよい。「極値付近」または「ほぼ極値をとる」としたのは、多少誤差を含んでいてもよいためである。多少の誤差は前提技術にはさして問題とならない。

## 【 0 0 3 4 】

極値自体も前記パラメータに依存するため、極値の挙動、つまり極値の変化の様子をもとに、最適と考えられるパラメータを決定する余地が生じる。この態様はその事実を利用している。この態様によれば、元来調整の困難なパラメータの決定を自動化する途が拓かれる。

## 【 0 0 3 5 】

## [前提技術の実施の形態]

最初に [ 1 ] で前提技術の要素技術の詳述し、[ 2 ] で処理手順を具体的に説明する。さらに [ 3 ] で実験の結果を報告する。

## 【 0 0 3 6 】

## [ 1 ] 要素技術の詳細

## 〔 1 . 1 〕 イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

## 【 0 0 3 7 】

本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリュームレンダリング、少ないフレームからの滑らかな動画像の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリュームレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

## 【 0 0 3 8 】

## 〔 1 . 2 〕 特異点フィルタの階層

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅を  $N$ 、高さを  $M$  とする。以下簡単のため、 $N = M = 2^n$  ( $n$  は自然数) と仮定する。また、区間  $[0, N] \subset \mathbb{R}$  を  $I$  と記述する。 $(i, j)$  における画像の画素を  $p(i, j)$  と記述する ( $i, j \in I$ )。

## 【 0 0 3 9 】

ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第  $m$  レベルにおける各画像のサイズは  $2^m \times 2^m$  ( $0 \leq m \leq n$ ) とする。特異点フィルタは次の 4 種類の新たな階層画像を  $n$  から下がる方向で再帰的に構築する。

## 【 0 0 4 0 】

【数 1】

$$\begin{aligned}
 p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\
 p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\
 p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\
 p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)}))
 \end{aligned}$$

(式 1)

ただしここで、

【0 0 4 1】

【数 2】

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式 2})$$

とする。以降これら 4 つの画像を副画像（サブイメージ）と呼ぶ。 $\min_{x \leq t} \leq x+1$ 、 $\max_{x \leq t \leq x+1}$  をそれぞれ  $\alpha$  及び  $\beta$  と記述すると、副画像はそれぞれ以下のように記述できる。

【0 0 4 2】

$$\begin{aligned}
 p(m, 0) &= \alpha(x) \alpha(y) p(m+1, 0) \\
 p(m, 1) &= \alpha(x) \beta(y) p(m+1, 1) \\
 p(m, 2) &= \beta(x) \alpha(y) p(m+1, 2) \\
 p(m, 3) &= \beta(x) \beta(y) p(m+1, 3)
 \end{aligned}$$

すなわち、これらは  $\alpha$  と  $\beta$  のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかなように、特異点フィルタはもとの画像について  $2 \times 2$  画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるととも

に、他方について最小画素値となる画素は鞍点として検出される。

#### 【0043】

特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\alpha(x)$   $\alpha(y)$  は極小点を保存し、 $\beta(x)$   $\beta(y)$  は極大点を保存し、 $\alpha(x)$   $\beta(y)$  及び  $\beta(x)$   $\alpha(y)$  は鞍点を保存する。

#### 【0044】

はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

#### 【0045】

この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれられていく。まず  $p(m, 0)$  を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p(m, 1)$  を用いて鞍点のマッチングがとられ、 $p(m, 2)$  を用いて他の鞍点のマッチングがとられる。そして最後に  $p(m, 3)$  を用いて極大点のマッチングがとられる。

#### 【0046】

図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像  $p(5, 0)$  を示している。同様に、図1(e)と図1(f)は  $p(5, 1)$ 、図1(g)と図1(h)は  $p(5, 2)$ 、図1(i)と図1(j)は  $p(5, 3)$  をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず  $p(5, 0)$  によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p(5, 1)$  によれば口が明確になる。口は横方向で輝度が低いためである。 $p(5, 2)$  によれば首の両側の縦線が明確になる。最後に、 $p(5, 3)$  によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

#### 【0047】

特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0048】

[1. 3] 画像間の写像の計算

始点画像の位置  $(i, j)$  の画素を  $p^{(n)}(i, j)$  と書き、同じく終点画像の位置  $(k, l)$  の画素を  $q^{(n)}(k, l)$  で記述する。 $i, j, k, l \in I$  とする。画像間の写像のエネルギー（後述）を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ  $p^{(m, 0)}$  と  $q^{(m, 0)}$  間の写像  $f^{(m, 0)} : p^{(m, 0)} \rightarrow q^{(m, 0)}$  が計算される。 $f^{(m, 0)}$  に基づき、最小エネルギーを持つ  $p^{(m, 1)}$ 、 $q^{(m, 1)}$  間の写像  $f^{(m, 1)}$  が計算される。この手続は、 $p^{(m, 3)}$  と  $q^{(m, 3)}$  の間の写像  $f^{(m, 3)}$  の計算が終了するまで続く。各写像  $f^{(m, i)}$  ( $i = 0, 1, 2, \dots$ ) を副写像と呼ぶことにする。 $f^{(m, i)}$  の計算の都合のために、 $i$  の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【0049】

【数3】

$$f^{(m, i)} : p^{(m, \sigma(i))} \rightarrow q^{(m, \sigma(i))} \quad (\text{式 3})$$

ここで  $\sigma(i) \in \{0, 1, 2, 3\}$  である。

【0050】

[1. 3. 1] 全単射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射かつ単射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全単射のデジタル版である。前提技術では、画素は格子点によって特定される。

【0051】

始点副画像（始点画像について設けられた副画像）から終点副画像（終点画像について設けられた副画像）への写像は、 $f^{(m, s)} : I/2^{n-m} \times I/2^{n-m} \rightarrow I/2^{n-m} \times I/2^{n-m}$  ( $s = 0, 1, \dots$ ) によって表される。ここで、 $f^{(m, s)}(i, j) = (k, l)$  は、始点画像の  $p^{(m, s)}(i, j)$  が終点画像の  $q^{(m, s)}(k, l)$  に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$  が成り立つとき画素  $q(k, l)$  を  $q_f(i, j)$  と記述する。

【0052】

前提技術で扱う画素（格子点）のようにデータが離散的な場合、全単射の定義は重要である。ここでは以下のように定義する（ $i, i', j, j', k, l$  は全て整数とする）。まず始めに、始点画像の平面において  $R$  によって表記される各正方形領域、

【0053】

【数4】

$$p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}$$

(式4)

を考える（ $i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$ ）。ここで  $R$  の各辺（エッジ）の方向を以下のように定める。

【0054】

【数5】

$$\overrightarrow{p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)}}, \overrightarrow{p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)}}, \overrightarrow{p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{p_{(i,j+1)}^{(m,s)} p_{(i,j)}^{(m,s)}}$$

(式5)

この正方形は写像  $f$  によって終点画像平面における四辺形に写像されなければならない。 $f^{(m, s)}(R)$  によって示される四辺形、

【0055】

【数6】

$$q_{f(i,j)}^{(m,s)} q_{f(i+1,j)}^{(m,s)} q_{f(i+1,j+1)}^{(m,s)} q_{f(i,j+1)}^{(m,s)}$$

(式6)

は、以下の全単射条件を満たす必要がある。

## 【0056】

1. 四辺形  $f^{(m, s)}(R)$  のエッジは互いに交差しない。
2.  $f^{(m, s)}(R)$  のエッジの方向は  $R$  のそれらに等しい（図2の場合、時計回り）。
3. 緩和条件として収縮写像（リトラクション：retractions）を許す。

## 【0057】

何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないためである。ここでは  $f^{(m, s)}(R)$  のひとつのエッジの長さが0、すなわち  $f^{(m, s)}(R)$  は三角形になってもよい。しかし、面積が0となるような図形、すなわち1点または1本の線分になってはならない。図2（R）がもとの四辺形の場合、図2（A）と図2（D）は全単射条件を満たすが、図2（B）、図2（C）、図2（E）は満たさない。

## 【0058】

実際のインプリメンテーションでは、写像が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ （ただし  $i = 0, i = 2^m - 1, j = 0, j = 2^m - 1$  の4本の線上）である。この条件を以下「付加条件」とも呼ぶ。

## 【0059】

[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト

写像  $f$  のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像  $f^{(m, s)}$  の点  $(i, j)$  におけるエネルギー  $C^{(m, s)}(i, j)$  は次式によって定まる。

## 【0060】

【数7】

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{(i,j)}^{(m,s)})|^2 \quad (\text{式7})$$

ここで、 $V(p^{(m,s)}(i,j))$  及び  $V(q^{(m,s)}(i,j))$  はそれぞれ画素  $p^{(m,s)}(i,j)$  及び  $q^{(m,s)}(i,j)$  の輝度である。 $f$  のトータルのエネルギー  $C^{(m,s)}$  は、マッチングを評価するひとつの評価式であり、つぎに示す  $C^{(m,s)}(i,j)$  の合計で定義できる。

【0061】

【数8】

$$C_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式8})$$

[1.3.2.2] 滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー  $D_f$  を導入する。このエネルギーは画素の輝度とは関係なく、 $p^{(m,s)}(i,j)$  および  $q^{(m,s)}(i,j)$  の位置によって決まる ( $i=0, \dots, 2^m-1, j=0, \dots, 2^m-1$ )。点  $(i,j)$  における写像  $f^{(m,s)}$  のエネルギー  $D^{(m,s)}(i,j)$  は次式で定義される。

【0062】

【数9】

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)} \quad (\text{式9})$$

ただし、係数パラメータ  $\eta$  は0以上の実数であり、また、

【0063】

【数10】

$$E_{0(i,j)}^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2 \quad (\text{式10})$$

【0064】

【数11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2 / 4 \quad (\text{式11})$$

とする。ここで、



【0065】

【数12】

$$\|(x, y)\| = \sqrt{x^2 + y^2}$$

(式12)

であり、 $i' < 0$  および  $j' < 0$  に対して  $f(i', j')$  は 0 と決める。 $E_0$  は  $(i, j)$  及び  $f(i, j)$  の距離で決まる。 $E_0$  は画素があまりにも離れた画素へ写影されることを防ぐ。ただし  $E_0$  は、後に別のエネルギー関数で置き換える。 $E_1$  は写像の滑らかさを保証する。 $E_1$  は、 $P(i, j)$  の変位とその隣接点の変位の間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー  $D_f$  は次式で定まる。

【0066】

【数13】

$$D_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式13})$$

[1. 3. 2. 3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は  $\lambda C^{(m,s)}_f + D^{(m,s)}_f$  で定義される。ここで係数パラメータ  $\lambda$  は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【0067】

【数14】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (\text{式14})$$

$\lambda = 0$  及び  $\eta = 0$  の場合、写像は単位写像になることに注意すべきである（すなわち、全ての  $i = 0, \dots, 2^m - 1$  及び  $j = 0, \dots, 2^m - 1$  に対して  $f^{(m,s)}(i, j) = (i, j)$  となる）。後述のごとく、本前提技術では最初に  $\lambda = 0$  及び  $\eta = 0$  の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の  $\lambda$  の位置を変えて  $C^{(m,s)}_f + \lambda D^{(m,s)}_f$  と定義したとすれば、 $\lambda = 0$  及び  $\eta = 0$  の場合に総合評価式が  $C^{(m,s)}_f$  だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけ

で対応づけられ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していてもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【0068】

オブティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オブティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

【0069】

[1. 3. 3] 多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像  $f_{min}$  を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0070】

まず、

【0071】

【数15】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式15})$$

が成り立つとき、 $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$  をそれぞれ  $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$  の parent と呼ぶことにする。[x] は x を越えない最大整数である。また  $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$  をそれぞれ  $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$  の child と呼ぶ。関数 parent

(i, j) は次式で定義される。

【0072】

【数16】

$$\text{parent}(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式16})$$

$p(m, s)(i, j)$  と  $q(m, s)(k, l)$  の間の写像  $f(m, s)$  は、エネルギー計算を行って最小になったものを見つけることで決定される。 $f(m, s)(i, j) = (k, l)$  の値は  $f(m-1, s)(m=1, 2, \dots, n)$  を用いることによって、以下のように決定される。まず、 $q(m, s)(k, l)$  は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【0073】

【数17】

$$q_{g(m,s)}^{(m,s)}(i-1,j-1) q_{g(m,s)}^{(m,s)}(i-1,j+1) q_{g(m,s)}^{(m,s)}(i+1,j+1) q_{g(m,s)}^{(m,s)}(i+1,j-1) \quad (\text{式17})$$

ただしここで、

【0074】

【数18】

$$g^{(m,s)}(i, j) = f^{(m-1,s)}(\text{parent}(i, j)) + f^{(m-1,s)}(\text{parent}((i, j) + (1, 1))) \quad (\text{式18})$$

である。こうして定めた四辺形を、以下  $p(m, s)(i, j)$  の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0075】

図3は以上の手順を示している。同図において、始点画像のA, B, C, Dの画素は、第  $m-1$  レベルにおいてそれぞれ終点画像のA', B', C', D'へ写影される。画素  $p(m, s)(i, j)$  は、相続四辺形A' B' C' D'の内部に存在する画素  $q(m, s)f(m)(i, j)$  へ写影されなければならない

。以上の配慮により、第 $m-1$ レベルの写像から第 $m$ レベルの写像への橋渡しがなされる。

【0076】

先に定義したエネルギー $E_0$ は、第 $m$ レベルにおける副写像 $f^{(m, 0)}$ を計算するために、次式に置き換える。

【0077】

【数19】

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式19})$$

また、副写像 $f^{(m, s)}$ を計算するためには次式を用いる。

【0078】

【数20】

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq i) \quad (\text{式20})$$

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式19は、 $f^{(m, s)}(i, j)$ と、第 $m-1$ レベルの画素の一部と考えた場合の $(i, j)$ が射影されるべき点の位置との距離を示している。

【0079】

仮に、相続四辺形 $A' B' C' D'$ の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、 $A' B' C' D'$ の境界線からの距離が $L$ （始めは $L=1$ ）である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを $f^{(m, s)}(i, j)$ の値として選択する。そのような点が発見されるか、または $L$ がその上限の $L^{(m)}_{\max}$ に到達するまで、 $L$ を大きくしていく。 $L^{(m)}_{\max}$ は各レベル $m$ に対して固定である。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f^{(m, s)}(i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第1及び第2条件を外す。

## 【0080】

多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル（粗いレベル）において小さいためである。

## 【0081】

## [1. 4] 最適なパラメータ値の自動決定

既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

## 【0082】

前提技術に係るシステムはふたつのパラメータ、 $\lambda$  及び  $\eta$  を含む。端的にいえば、 $\lambda$  は画素の輝度の差の重みであり、 $\eta$  は写像の剛性を示している。これらのパラメータの値は初期値が 0 であり、まず  $\eta = 0$  に固定して  $\lambda$  を 0 から徐々に増加させる。 $\lambda$  の値を大きくしながら、しかも総合評価式（式 14）の値を最小にする場合、各副写像に関する  $C^{(m, s)}_f$  の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 $\lambda$  が最適値を超えると以下の現象が発生する。

## 【0083】

1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

## 【0084】

3. その結果、式14において $D(m, s)_f$ が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするため、 $D(m, s)_f$ の急激な増加を抑制するよう $f(m, s)$ が変化し、その結果 $C(m, s)_f$ が増加する。

したがって、 $\lambda$ を増加させながら式14が最小値をとるという状態を維持しつつ $C(m, s)_f$ が減少から増加に転じる閾値を検出し、その $\lambda$ を $\eta = 0$ における最適値とする。つぎに $\eta$ を少しずつ増やして $C(m, s)_f$ の挙動を検査し、後述の方法で $\eta$ を自動決定する。その $\eta$ に対応して $\lambda$ も決まる。

## 【0085】

この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できる時、その目が固定される。

## 【0086】

[1. 4. 1]  $\lambda$ の動的決定

$\lambda$ は0から所定の刻み幅で増加されていき、 $\lambda$ の値が変わる度に副写像が評価される。式14のごとく、総エネルギーは $\lambda C(m, s)_f + D(m, s)_f$ によって定義される。式9の $D(m, s)_f$ は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど $E_0$ も $E_1$ も増加していく。 $E_1$ は整数であるから、 $D(m, s)_f$ の最小刻み幅は1である。このため、現在の $\lambda C(m, s)_{(i, j)}$ の変化(減少量)が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って $D(m, s)_f$ は1以上増加するため、 $\lambda C(m, s)_{(i, j)}$ が1以上減少しない限り総エネルギーは減らないためである。

## 【0087】

この条件のもと、 $\lambda$ の増加に伴い、正常な場合に $C(m, s)_{(i, j)}$ が減少することを示す。 $C(m, s)_{(i, j)}$ のヒストグラムを $h(1)$ と記述する。 $h(1)$ はエネルギー $C(m, s)_{(i, j)}$ が $1^2$ である画素の数である。 $\lambda 1^2 \geq 1$ が成り立つために、例えば $1^2 = 1/\lambda$ の場合を考える。 $\lambda$ が $\lambda_1$

から  $\lambda_2$  まで微小量変化するとき、

【0088】

【数21】

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lfloor \frac{1}{\lambda_1} \rfloor} h(l) \simeq \int_{l=\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式21)

で示されるA個の画素が、

【0089】

【数22】

$$C_f^{(m,s)} - l^2 = C_f^{(m,s)} - \frac{1}{\lambda} \quad (\text{式22})$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロになると近似している。この式は  $C^{(m,s)}_f$  の値が、

【0090】

【数23】

$$\partial C_f^{(m,s)} = -\frac{A}{\lambda} \quad (\text{式23})$$

だけ変化することを示し、その結果、

【0091】

【数24】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (\text{式24})$$

が成立する。 $h(l) > 0$  であるから、通常  $C^{(m,s)}_f$  は減少する。しかし、 $\lambda$  が最適値を越えようとするとき、上述の現象、つまり  $C^{(m,s)}_f$  の増加が発生する。この現象を検出することにより、 $\lambda$  の最適値を決定する。

【0092】

なお、 $H(h > 0)$  及び  $k$  を定数とするとき、

【0093】

【数 2 5】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (\text{式 2 5})$$

と仮定すれば、

【0 0 9 4】

【数 2 6】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式 2 6})$$

が成り立つ。このとき  $k \neq -3$  であれば、

【0 0 9 5】

【数 2 7】

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (\text{式 2 7})$$

となる。これが  $C^{(m, s)}_f$  の一般式である ( $C$  は定数)。

【0 0 9 6】

$\lambda$  の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を  $p_0$  と仮定する。この場合、

【0 0 9 7】

【数 2 8】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式 2 8})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

【0 0 9 8】

【数 2 9】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}} \quad (\text{式 2 9})$$

従って、

【0 0 9 9】



【数30】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (\text{式30})$$

は定数である。仮に  $h(1) = H 1^k$  を仮定するとき、例えば、

【0100】

【数31】

$$B_0 \lambda^{3/2+k/2} = p_0 H \quad (\text{式31})$$

は定数になる。しかし  $\lambda$  が最適値を越えると、上の値は急速に増加する。この現象を検出し、 $B_0 \lambda^{3/2+k/2} / 2^m$  の値が異常値  $B_{0 \text{ thres}}$  を越えるかどうかを検査し、 $\lambda$  の最適値を決定することができる。同様に、 $B_1 \lambda^{3/2+k/2} / 2^m$  の値が異常値  $B_{1 \text{ thres}}$  を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率  $B_1$  を確認する。ファクター  $2^m$  を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギー  $C(m, s)_f$  の観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0101】

なお実験では、副写像  $f(m, s)$  を計算する際、もし  $\lambda$  が0.1を越えたら  $f(m, s)$  の計算は止めて  $f(m, s+1)$  の計算に移行した。 $\lambda > 0.1$  のとき、画素の輝度255レベル中のわずか「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0.1$  のとき正しい結果を得ることは困難だったためである。

【0102】

[1.4.2] ヒストグラム  $h(1)$ 

$C(m, s)_f$  の検査はヒストグラム  $h(1)$  に依存しない。全単射及びその第3の条件の検査の際、 $h(1)$  に影響を受けうる。実際に  $(\lambda, C(m, s)_f)$  をプロットすると、 $k$  は通常1付近にある。実験では  $k=1$  を用い、 $B_0 \lambda^2$  と  $B_1 \lambda^2$  を検査した。仮に  $k$  の本当の値が1未満であれば、 $B_0 \lambda^2$  と  $B_1 \lambda^2$  は定数にならず、ファクター  $\lambda^{(1-k)/2}$  に従って徐々に増加する。 $h(1)$  が定数であれば、例えばファクターは  $\lambda^{1/2}$  である。しかし、こうした

差は閾値  $B_{0thres}$  を正しく設定することによって吸収することができる。

【0103】

ここで次式のごとく始点画像を中心が  $(x_0, y_0)$ 、半径  $r$  の円形のオブジェクトであると仮定する。

【0104】

【数32】

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (otherwise) \end{cases} \quad (式32)$$

一方、終点画像は、次式のごとく中心  $(x_1, y_1)$ 、半径が  $r$  のオブジェクトであるとする。

【0105】

【数33】

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (otherwise) \end{cases} \quad (式33)$$

ここで  $c(x)$  は  $c(x) = x^k$  の形であるとする。中心  $(x_0, y_0)$  及び  $(x_1, y_1)$  が十分遠い場合、ヒストグラム  $h(l)$  は次式の形となる。

【0106】

【数34】

$$h(l) \propto r l^k \quad (k \neq 0) \quad (式34)$$

$k=1$  のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k=-1$  のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般のオブジェクトはこれらふたつのタイプのオブジェクトの間にあると考えてもさして一般性を失わない。したがって、 $k$  は  $-1 \leq k \leq 1$  として大抵の場合をカバーでき、式27が一般に減少関数であることが保障される。

【0107】

なお、式34からわかるように、 $r$  は画像の解像度に影響されること、すなわち  $r$  は  $2^m$  に比例することに注意すべきである。このために [1. 4. 1] にお

いてファクター  $2^m$  を導入した。

【0108】

#### [1. 4. 3] $\eta$ の動的決定

パラメータ  $\eta$  も同様の方法で自動決定できる。はじめに  $\eta = 0$  とし、最も細かい解像度における最終的な写像  $f^{(n)}$  及びエネルギー  $C_f^{(n)}$  を計算する。つづいて、 $\eta$  をある値  $\Delta \eta$  だけ増加させ、再び最も細かい解像度における最終写像  $f^{(n)}$  及びエネルギー  $C_f^{(n)}$  を計算し直す。この過程を最適値が求まるまで続ける。 $\eta$  は写像の剛性を示す。次式の重みだからである。

【0109】

【数35】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式35)

$\eta$  が 0 のとき、 $D_f^{(n)}$  は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 $\eta$  が非常に大きな値のとき、 $D_f^{(n)}$  は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 $\eta$  の値が 0 から次第に増えるとき、後述のごとく  $C_f^{(n)}$  は徐々に減少する。しかし  $\eta$  の値が最適値を越えると、図 4 に示すとおり、エネルギーは増加し始める。同図の X 軸は  $\eta$ 、Y 軸は  $C_f$  である。

【0110】

この方法で  $C_f^{(n)}$  を最小にする最適な  $\eta$  の値を得ることができる。しかし、 $\lambda$  の場合に比べていろいろな要素が計算に影響する結果、 $C_f^{(n)}$  は小さく揺らぎながら変化する。 $\lambda$  の場合は、入力が微小量変化するたびに副写像を 1 回計算しなおすだけだが、 $\eta$  の場合はすべての副写像が計算しなおされるためである。このため、得られた  $C_f^{(n)}$  の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つければ、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0111】

## [ 1 . 5 ] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f(m, s)$  の値域を  $R \times R$  に拡張することができる ( $R$  は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

【 0 1 1 2 】

【数 3 6】

$$V(q_{f(m,s)}^{(m,s)}(i,j)) \quad (式 3 6)$$

における輝度を持つ  $f(m, s)$  が提供される。つまりスーパーサンプリングが行われる。実験では、 $f(m, s)$  は整数及び半整数値をとることが許され、

【 0 1 1 3 】

【数 3 7】

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (式 3 7)$$

は、

【 0 1 1 4 】

【数 3 8】

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (式 3 8)$$

によって与えられた。

【 0 1 1 5 】

## [ 1 . 6 ] 各画像の画素の輝度の正規化

始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー  $C(m, s)_f$  が大きくなりすぎ、正しい評価がしづらいためである。

【 0 1 1 6 】

例えば、人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふた

つの顔の間の副写像を計算するために、まず副画像を正規化する。すなわち、最も暗い画素の輝度を0、最も明るいそれを255に設定し、他の画素の輝度は線形補間によって求めておく。

【0117】

#### [1. 7] インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1番上の左端の画素  $(i, j) = (0, 0)$  について  $f(m, s)$  の値を決定する。次に  $i$  を1ずつ増やしながら各  $f(m, s)(i, j)$  の値を決定する。 $i$  の値が画像の幅に到達したとき、 $j$  の値を1増やし、 $i$  を0に戻す。以降、始点画像のスキャンに伴い  $f(m, s)(i, j)$  を決定していく。すべての点について画素の対応が決まれば、ひとつの写像  $f(m, s)$  が決まる。

【0118】

ある  $p(i, j)$  について対応点  $q_f(i, j)$  が決まれば、つぎに  $p(i, j+1)$  の対応点  $q_f(i, j+1)$  が決められる。この際、 $q_f(i, j+1)$  の位置は全単射条件を満たすために、 $q_f(i, j)$  の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに  $(0, 0)$  が最も優先される状態がつづく、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f(m, s)$  を以下の方法で決めていく。

【0119】

まず  $(s \bmod 4)$  が0の場合、 $(0, 0)$  を開始点とし  $i$  及び  $j$  を徐々に増やしながら決めていく。 $(s \bmod 4)$  が1の場合、最上行の右端点を開始点とし、 $i$  を減少、 $j$  を増加させながら決めていく。 $(s \bmod 4)$  が2のとき、最下行の右端点を開始点とし、 $i$  及び  $j$  を減少させながら決めていく。 $(s \bmod 4)$  が3の場合、最下行の左端点を開始点とし、 $i$  を増加、 $j$  を減少させながら決めていく。解像度が最も細かい第  $n$  レベルには副写像という概念、すなわちパラメータ  $s$  が存在しないため、仮に  $s = 0$  及び  $s = 2$  であるとしてふたつの方向を連続的に計算した。

【0120】

実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補  $(k, 1)$  の中からできる限り全単射条件を満たす  $f(m, s)(i, j)$  ( $m=0, \dots, n$ ) の値を選んだ。第3の条件を破る候補のエネルギー  $D(k, 1)$  には  $\phi$  を掛け、一方、第1または第2の条件を破る候補には  $\psi$  を掛ける。今回は  $\phi=2$ 、 $\psi=100000$  を用いた。

【0121】

前述の全単射条件のチェックのために、実際の手続として  $(k, 1) = f(m, s)(i, j)$  を決定する際に以下のテストを行った。すなわち  $f(m, s)(i, j)$  の相続四辺形に含まれる各格子点  $(k, 1)$  に対し、次式の外積の  $z$  成分が0以上になるかどうかを確かめる。

【0122】

【数39】

$$W = \vec{A} \times \vec{B} \quad (\text{式39})$$

ただしここで、

【0123】

【数40】

$$\vec{A} = \overrightarrow{q_{f(m,s)(i,j-1)}^{(m,s)} q_{f(m,s)(i+1,j-1)}^{(m,s)}} \quad (\text{式40})$$

【0124】

【数41】

$$\vec{B} = \overrightarrow{q_{f(m,s)(i,j-1)}^{(m,s)} q_{(k,l)}^{(m,s)}} \quad (\text{式41})$$

である（ここでベクトルは三次元ベクトルとし、 $z$  軸は直交右手座標系において定義される）。もし  $W$  が負であれば、その候補については  $D(m, s)(k, 1)$  に  $\phi$  を掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0125】

図5(a)、図5(b)はこの条件を検査する理由を示している。図5(a)

はペナルティのない候補、図5 (b) はペナルティがある候補をそれぞれ表す。隣接画素  $(i, j+1)$  に対する写像  $f^{(m, s)}(i, j+1)$  を決定する際、 $W$  の  $z$  成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、 $q^{(m, s)}(k, 1)$  は隣接する四辺形の境界線を越えるためである。

## 【0126】

## [1. 7. 1] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときには  $\sigma(0) = 0$ 、 $\sigma(1) = 1$ 、 $\sigma(2) = 2$ 、 $\sigma(3) = 3$ 、 $\sigma(4) = 0$  を用い、奇数のときは  $\sigma(0) = 3$ 、 $\sigma(1) = 2$ 、 $\sigma(2) = 1$ 、 $\sigma(3) = 0$ 、 $\sigma(4) = 3$  を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、 $s$  は0～3のいずれかである。しかし、実際には  $s = 4$  に相当する処理を行った。その理由は後述する。

## 【0127】

## [1. 8] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形  $p(i, j) p(i+1, j) p(i, j+1) p(i+1, j+1)$  が終点画像平面上の四辺形  $q f(i, j) q f(i+1, j) q f(i, j+1) q f(i+1, j+1)$  に射影されると仮定する。簡単のため、画像間の距離を1とする。始点画像平面からの距離が  $t$  ( $0 \leq t \leq 1$ ) である中間画像の画素  $r(x, y, t)$  ( $0 \leq x \leq N-1$ ,  $0 \leq y \leq M-1$ ) は以下の要領で求められる。まず画素  $r(x, y, t)$  の位置 (ただし  $x, y, t \in \mathbb{R}$ ) を次式で求める。

## 【0128】

【数 4 2】

$$\begin{aligned}
 (x, y) = & (1 - dx)(1 - dy)(1 - t)(i, j) + (1 - dx)(1 - dy)tf(i, j) \\
 & + dx(1 - dy)(1 - t)(i + 1, j) + dx(1 - dy)tf(i + 1, j) \\
 & + (1 - dx)dy(1 - t)(i, j + 1) + (1 - dx)dytf(i, j + 1) \\
 & + dxdy(1 - t)(i + 1, j + 1) + dxdytf(i + 1, j + 1)
 \end{aligned}$$

(式 4 2)

つづいて  $r(x, y, t)$  における画素の輝度が次の式を用いて決定される。

【0 1 2 9】

【数 4 3】

$$\begin{aligned}
 V(r(x, y, t)) = & (1 - dx)(1 - dy)(1 - t)V(p_{(i, j)}) + (1 - dx)(1 - dy)tV(q_{f(i, j)}) \\
 & + dx(1 - dy)(1 - t)V(p_{(i+1, j)}) + dx(1 - dy)tV(q_{f(i+1, j)}) \\
 & + (1 - dx)dy(1 - t)V(p_{(i, j+1)}) + (1 - dx)dytV(q_{f(i, j+1)}) \\
 & + dxdy(1 - t)V(p_{(i+1, j+1)}) + dxdytf(q_{f(i+1, j+1)})
 \end{aligned}$$

(式 4 3)

ここで  $dx$  及び  $dy$  はパラメータであり、0 から 1 まで変化する。

【0 1 3 0】

【1. 9】拘束条件を課したときの写像

いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0 1 3 1】

基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像  $f$  を正確に計算する。

【0 1 3 2】

まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第  $m$  レベルの大まかな写像を  $F^{(m)}$  と記述する。



【 0 1 3 3 】

大まかな写像 F は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について  $n_s$  個の画素、

【 0 1 3 4 】

【数 4 4】

$$P(i_0, j_0), P(i_1, j_1), \dots, P(i_{n_s-1}, j_{n_s-1}) \quad (\text{式 4 4})$$

を特定するとき、以下の値を決める。

【 0 1 3 5 】

【数 4 5】

$$\begin{aligned} F^{(n)}(i_0, j_0) &= (k_0, l_0), \\ F^{(n)}(i_1, j_1) &= (k_1, l_1), \dots, \\ F^{(n)}(i_{n_s-1}, j_{n_s-1}) &= (k_{n_s-1}, l_{n_s-1}) \end{aligned}$$

(式 4 5)

始点画像の他の画素の変位量は、 $P(i_h, j_h)$  ( $h = 0, \dots, n_s - 1$ ) の変位に重み付けをして求められる平均である。すなわち画素  $P(i, j)$  は、終点画像の以下の画素に射影される。

【 0 1 3 6 】

【数 4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{n_s-1} (k_h - i, l_h - j) \text{weight}_h(i, j)}{2^{n-m}} \quad (\text{式 4 6})$$

ただしここで、

【 0 1 3 7 】

【数 4 7】

$$\text{weight}_h(i, j) = \frac{1/||i_h - i, j_h - j||^2}{\text{total weight}(i, j)} \quad (\text{式 4 7})$$

【 0 1 3 8 】

【数 4 8】

$$total\ weight(i, j) = \sum_{h=0}^{h=n_s-1} 1/\|(i_h - i, j_h - j)\|^2 \quad (式 4 8)$$

とする。

【0 1 3 9】

つづいて、 $F^{(m)}$  に近い候補写像  $f$  がより少ないエネルギーを持つように、その写像  $f$  のエネルギー  $D^{(m, s)}(i, j)$  を変更する。正確には、 $D^{(m, s)}(i, j)$  は、

【0 1 4 0】

【数 4 9】

$$D_{(i,j)}^{(m,s)} = E_{0(i,j)}^{(m,s)} + \eta E_{1(i,j)}^{(m,s)} + \kappa E_{2(i,j)}^{(m,s)} \quad (式 4 9)$$

である。ただし、

【0 1 4 1】

【数 5 0】

$$E_{2(i,j)}^{(m,s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i, j) - f^{(m,s)}(i, j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i, j) - f^{(m,s)}(i, j)\|^2, & \text{otherwise} \end{cases} \quad (式 5 0)$$

であり、 $\kappa, \rho \geq 0$  とする。最後に、前述の写像の自動計算プロセスにより、 $f$  を完全に決定する。

【0 1 4 2】

ここで、 $f^{(m, s)}(i, j)$  が  $F^{(m)}(i, j)$  に十分近いとき、つまりそれらの距離が、

【0 1 4 3】

【数 5 1】

$$\left\lfloor \frac{\rho^2}{2^{2(n-m)}} \right\rfloor \quad (式 5 1)$$

以内であるとき、 $E_2^{(m, s)}(i, j)$  が 0 になることに注意すべきである

。そのように定義した理由は、各  $f^{(m, s)}(i, j)$  が  $F^{(m)}(i, j)$  に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

## 【 0 1 4 4 】

## 〔 2 〕 具体的な処理手順

〔 1 〕 の各要素技術による処理の流れを説明する。

図 6 は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S 1)、つづいて始点画像と終点画像のマッチングをとる (S 2)。ただし、S 2 は必須ではなく、S 1 で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

## 【 0 1 4 5 】

図 7 は図 6 の S 1 の詳細を示すフローチャートである。ここでは S 2 で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い (S 1 0)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い (S 1 1)、一連の終点階層画像を得る。ただし、S 1 0 と S 1 1 の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

## 【 0 1 4 6 】

図 8 は図 7 の S 1 0 の詳細を示すフローチャートである。もとの始点画像のサイズは  $2^n \times 2^n$  とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ  $m$  を  $n$  にセットする (S 1 0 0)。つづいて第  $m$  レベルの画像  $p^{(m, 0)}$ 、 $p^{(m, 1)}$ 、 $p^{(m, 2)}$ 、 $p^{(m, 3)}$  から特異点フィルタを用いて特異点を検出し (S 1 0 1)、それぞれ第  $m-1$  レベルの画像  $p^{(m-1, 0)}$ 、 $p^{(m-1, 1)}$ 、 $p^{(m-1, 2)}$ 、 $p^{(m-1, 3)}$  を生成する (S 1 0 2)。ここでは  $m=n$  であるため、 $p^{(m, 0)} = p^{(m, 1)} = p^{(m, 2)} = p^{(m, 3)} = p^{(n)}$  であり、ひとつの始点画像から 4 種類の副画像が生成される。

## 【 0 1 4 7 】

図9は第 $m$ レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p(m, s)$ は $p(m, 0) \sim p(m, 3)$ の4つの画像を象徴するもので、 $p(m-1, 0)$ を生成する場合には、 $p(m, s)$ は $p(m, 0)$ であると考え、[1. 2]で示した規則により、 $p(m-1, 0)$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p(m-1, 1)$ は「8」、 $p(m-1, 2)$ は「6」、 $p(m-1, 3)$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

## 【0148】

つづいて $m$ をデクリメントし(図8のS103)、 $m$ が負になっていないことを確認し(S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点でS10が終了する。第0レベルの副画像のサイズは $1 \times 1$ である。

## 【0149】

図10はS10によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7のS11にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6のS1による処理が完了する。

## 【0150】

前提技術では、図6のS2に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される(S30)。[1. 3. 2. 1]で導入した画素に関するエネルギー $C(m, s)_f$ と[1. 3. 2. 2]で導入した写像の滑らかさに関するエネルギー $D(m, s)_f$ がそれぞれである。つぎに、これらの評価式を統合して総合評価式を立てる(S31)。[1. 3. 2. 3]で導入した総エネルギー $\lambda C(m, s)_f + D(m, s)_f$ がそれぞれであり、[1. 3. 2. 2]で導入した $\eta$ を用いれば、

$$\sum \sum (\lambda C(m, s)(i, j) + \eta E_0(m, s)(i, j) + E_1(m$$

,  $s$ )  $(i, j)$  (式 5 2)

となる。ただし、総和は  $i, j$  についてそれぞれ  $0, 1 \dots, 2^m - 1$  で計算する。  
。以上でマッチング評価の準備が整う。

#### 【 0 1 5 1 】

図 1 2 は図 6 の S 2 の詳細を示すフローチャートである。[ 1 ] で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

#### 【 0 1 5 2 】

図 1 2 のごとく、まず係数パラメータ  $\eta$  を 0、レベルパラメータ  $m$  を 0 に設定する (S 2 0)。つづいて、始点階層画像中の第  $m$  レベルの 4 つの副画像と終点階層画像中の第  $m$  レベルの 4 つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような 4 種類の副写像  $f(m, s)$  ( $s = 0, 1, 2, 3$ ) を求める (S 2 1)。全単射条件は [ 1 . 3 . 3 ] で述べた相続四辺形を用いて検査される。この際、式 1 7、1 8 が示すように、第  $m$  レベルにおける副写像は第  $m - 1$  レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま  $m = 0$  であってそれより粗いレベルはないが、この例外的な処理は図 1 3 で後述する。

#### 【 0 1 5 3 】

一方、同一レベル内における水平的参照も行われる。[ 1 . 3 . 3 ] の式 2 0 のごとく、 $f(m, 3)$  は  $f(m, 2)$  に、 $f(m, 2)$  は  $f(m, 1)$  に、 $f(m, 1)$  は  $f(m, 0)$  に、それぞれ類似するように決める。その理由は、特異点の種類が違って、それらがもともとと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式 2 0 からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが

良好とみなされる。

#### 【0154】

なお、最初に決めるべき  $f(m, 0)$  については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験では  $f(m, 3)$  まで求まった後、これを拘束条件として  $f(m, 0)$  を一回更新するという手続をとった。これは式20に  $s = 4$  を代入し、 $f(m, 4)$  を新たな  $f(m, 0)$  とすることに等しい。 $f(m, 0)$  と  $f(m, 3)$  の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では[1. 7. 1]に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、 $s$  の値にしたがって開始点の位置を変える点は[1. 7]で述べたとおりである。

#### 【0155】

図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像  $f(0, s)$  はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、 $p(1, s)$  の点  $x$  の対応点を  $q(1, s)$  の中で探すとき、以下の手順を踏む。

#### 【0156】

1. 第1レベルの解像度で点  $x$  の左上点  $a$ 、右上点  $b$ 、左下点  $c$ 、右下点  $d$  を求める。

2. 点  $a \sim d$  がひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点  $a \sim d$  はそれぞれ画素  $A \sim D$  に属する。ただし、画素  $A \sim C$  は本来存在しない仮想的な画素である。

3. 第0レベルですでに求まっている画素  $A \sim D$  の対応点  $A' \sim D'$  を  $q(1, s)$  の中にプロットする。画素  $A' \sim C'$  は仮想的な画素であり、それぞれ画素  $A \sim C$  と同じ位置にあるものとする。

4. 画素  $A$  の中の点  $a$  の対応点  $a'$  が画素  $A'$  の中にあるとみなし、点  $a'$  を

プロットする。このとき、点  $a$  が画素  $A$  の中で占める位置（この場合、右下）と、点  $a'$  が画素  $A'$  の中で占める位置が同じであると仮定する。

5. 4と同様の方法で対応点  $b' \sim d'$  をプロットし、点  $a' \sim d'$  で相続四辺形を作る。

6. 相続四辺形の中でエネルギーが最小になるよう、点  $x$  の対応点  $x'$  を探す。対応点  $x'$  の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図 1 4 の場合、4 つの画素がすべて候補になる。

【0 1 5 7】

以上がある点  $x$  の対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第 2 レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図 3 に示すように画素  $A' \sim D'$  の間隔が空いていく状況が発生する。

【0 1 5 8】

こうして、ある第  $m$  レベルの 4 つの副写像が決まれば、 $m$  をインクリメントし（図 1 2 の S 2 2）、 $m$  が  $n$  を超えていないことを確かめて（S 2 3）、S 2 1に戻る。以下、S 2 1に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後に S 2 1に戻ったときに第  $n$  レベルの写像  $f^{(n)}$  を決める。この写像は  $\eta = 0$  に関して定まったものであるから、 $f^{(n)} (\eta = 0)$  と書く。

【0 1 5 9】

つぎに異なる  $\eta$  に関する写像も求めるべく、 $\eta$  を  $\Delta \eta$  だけシフトし、 $m$  をゼロクリアする（S 2 4）。新たな  $\eta$  が所定の探索打ち切り値  $\eta_{\max}$  を超えていないことを確認し（S 2 5）、S 2 1に戻り、今回の  $\eta$  に関して写像  $f^{(n)} (\eta = \Delta \eta)$  を求める。この処理を繰り返し、S 2 1で  $f^{(n)} (\eta = i \Delta \eta)$  ( $i = 0, 1, \dots$ ) を求めていく。 $\eta$  が  $\eta_{\max}$  を超えたとき S 2 6に進み、後述の方法で最適な  $\eta = \eta_{\text{opt}}$  を決定し、 $f^{(n)} (\eta = \eta_{\text{opt}})$  を最終的に写像  $f^{(n)}$  とする。

【0 1 6 0】

図 1 5 は図 1 2 の S 2 1 の詳細を示すフローチャートである。このフローチャートにより、ある定まった  $\eta$  について、第  $m$  レベルにおける副写像が決まる。副

写像を決める際、前提技術では副写像ごとに最適な $\lambda$ を独立して決める。

#### 【0161】

同図のごとく、まず $s$ と $\lambda$ をゼロクリアする(S210)。つぎに、そのときの $\lambda$ について(および暗に $\eta$ について)エネルギーを最小にする副写像 $f^{(m, s)}$ を求め(S211)、これを $f^{(m, s)} (\lambda=0)$ と書く。異なる $\lambda$ に関する写像も求めるべく、 $\lambda$ を $\Delta\lambda$ だけシフトし、新たな $\lambda$ が所定の探索打ち切り値 $\lambda_{max}$ を超えていないことを確認し(S213)、S211に戻り、以降の繰り返し処理で $f^{(m, s)} (\lambda=i\Delta\lambda) (i=0, 1, \dots)$ を求める。 $\lambda$ が $\lambda_{max}$ を超えたときS214に進み、最適な $\lambda=\lambda_{opt}$ を決定し、 $f^{(m, s)} (\lambda=\lambda_{opt})$ を最終的に写像 $f^{(m, s)}$ とする(S214)。

#### 【0162】

つぎに、同一レベルにおける他の副写像を求めるべく、 $\lambda$ をゼロクリアし、 $s$ をインクリメントする(S215)。 $s$ が4を超えていないことを確認し(S216)、S211に戻る。 $s=4$ になれば上述のごとく $f^{(m, 3)}$ を利用して $f^{(m, 0)}$ を更新し、そのレベルにおける副写像の決定を終了する。

#### 【0163】

図16は、ある $m$ と $s$ について $\lambda$ を変えながら求められた $f^{(m, s)} (\lambda=i\Delta\lambda) (i=0, 1, \dots)$ に対応するエネルギー $C^{(m, s)}_f$ の挙動を示す図である。[1.4]で述べたとおり、 $\lambda$ が増加すると通常 $C^{(m, s)}_f$ は減少する。しかし、 $\lambda$ が最適値を超えると $C^{(m, s)}_f$ は増加に転じる。そこで本前提技術では $C^{(m, s)}_f$ が極小値をとるときの $\lambda$ を $\lambda_{opt}$ と決める。同図のように $\lambda>\lambda_{opt}$ の範囲で再度 $C^{(m, s)}_f$ が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 $\lambda_{opt}$ は副写像ごとに独立して決めていき、最後に $f^{(n)}$ についてもひとつ定まる。

#### 【0164】

一方、図17は、 $\eta$ を変えながら求められた $f^{(n)} (\eta=i\Delta\eta) (i=0, 1, \dots)$ に対応するエネルギー $C^{(n)}_f$ の挙動を示す図である。ここでも $\eta$ が増加すると通常 $C^{(n)}_f$ は減少するが、 $\eta$ が最適値を超えると $C^{(n)}_f$ は



増加に転じる。そこで  $C^{(n)}_f$  が極小値をとるときの  $\eta$  を  $\eta_{opt}$  と決める。図 1 7 は図 4 の横軸のゼロ付近を拡大した図と考えてよい。 $\eta_{opt}$  が決まれば  $f^{(n)}$  を最終決定することができる。

## 【0 1 6 5】

以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先験的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

## 【0 1 6 6】

なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の 2 枚の画像間のマッチングをとる場合全般に利用できる。

## 【0 1 6 7】

たとえば 2 枚の画像間で、画素の輝度の差に関するエネルギー  $E_0$  と画素の位置的なずれに関するエネルギー  $E_1$  のふたつを評価式とし、これらの線形和  $E_{tot} = \alpha E_0 + E_1$  を総合評価式とする。この総合評価式の極値付近に注目して  $\alpha$  を自動決定する。つまり、いろいろな  $\alpha$  について  $E_{tot}$  が最小になるような写像を求める。それらの写像のうち、 $\alpha$  に関して  $E_1$  が極小値をとるときの  $\alpha$  を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

## 【0 1 6 8】

これ以外にも評価式の設定にはいろいろな方法があり、例えば  $1/E_1$  と  $1/E_2$  のように、評価結果が良好なほど大きな値をとるものを採用してもよい。総合評価式も必ずしも線形和である必要はなく、 $n$  乗和 ( $n=2$ 、 $1/2$ 、 $-1$ 、 $-2$  など)、多項式、任意の関数などを適宜選択すればよい。

## 【0169】

パラメータも、 $\alpha$ のみ、前提技術のごとく $\eta$ と $\lambda$ のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

## 【0170】

(2) 本前提技術では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C(m, s)_f$ が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば $\alpha E_0 + \beta E_1$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

## 【0171】

(3) 前提技術では各解像度レベルで4種類の特異点に関する4種類の副画像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f(m, 3)$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 $s$ に関する計算量が減る効果がある。

## 【0172】

(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が1/4になった。例えば $3 \times 3$ で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は1/9になる。

## 【0173】

(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像に変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

## 【0174】

### 〔画像補間に関する実施の形態〕

以上の前提技術を利用した画像補間技術を説明する。この技術によれば、従来は $1^{\circ}$ 程度ごとに撮影が必要だった商品の回転プレゼンテーションが、 $10 \sim 30^{\circ}$ の間隔でよいことが実験で確認されている。すなわち、この用途だけでも従来の $1/10 \sim 1/30$ のデータ量で同等以上のプレゼンテーションが可能である。

図18は、実施の形態に係る技術の概要を説明する図である。図18(a)、図18(b)、図18(c)、図18(d)は異なる視点からコーヒーカップを撮影した4枚のキーフレームである。キーフレームは予め撮影され準備された実在の画像であり、実施の形態の目的は、これらのキーフレームから図18(e)に示す中間フレームを生成することにある。図18(a)のキーフレームと図18(b)のキーフレームが第1画像対を形成し、図18(c)のキーフレームと図18(d)のキーフレームが第2画像対を形成する。目的の中間フレームは、第1画像対、第2画像対のいずれか一方では正しく生成できず、両方が揃って生成できる。中間フレームは、複数のキーフレームを縦横二方向について補間することで生成されるためである。

#### 【0175】

実施の形態によってユーザの視点に応じた中間フレームを連続的に生成することにより、少ないキーフレームから対象物の回転が等、擬似的な三次元画像を実現することができる。この特性は、ECにおける商品プレゼンテーション、動画像の圧縮、映像効果等に有用である。

#### 【0176】

図19は実施の形態に係る画像補間装置10の構成図である。画像補間装置10は、ユーザと相互作用をするGUI（グラフィカル・ユーザ・インタフェース）12と、GUI12を介して、生成すべき中間フレームの位置情報28を取得する中間フレーム位置取得部14と、複数のキーフレームを予め記憶しているキーフレーム記憶部16と、位置情報28をもとに中間フレームを生成するために必要なキーフレームをキーフレーム記憶部16から選定し、前提技術に基づいてマッチング計算を行うマッチングプロセッサ18と、その結果得られたキーフレ

ーム間の対応点情報を対応点ファイルとして記録する対応点ファイル保存部 2 0 と、対応点ファイルおよび前記位置情報 2 8 をもとに中間フレームを補間計算で生成する中間フレーム生成部 2 2 を含む。この装置 1 0 はさらに、生成された中間フレームをユーザの指示または視点に応じて連続的に表示する表示部 2 4 と、マッチングプロセッサ 1 8 によって計算された対応点ファイルを必要に応じて外部へ送出する通信部 2 6 を含む。

## 【 0 1 7 7 】

図 2 0 は、空間的に分散されたキーフレームと中間フレームの位置関係を示す。キーフレームはそれらを撮影した視点位置にしたがって並べられており、生成すべき中間フレームもその仮想的な視点位置に応じて同図の中に位置決めされる。

## 【 0 1 7 8 】

ここでは、9 枚のキーフレームである第 1 キーフレーム I 1 ～第 9 キーフレーム I 9 が準備されている。いま G U I 1 2 を通し、生成すべき中間フレームの位置が同図の中間フレーム I c と取得されたとき、まずその中間フレーム I c を取り巻くキーフレーム（以下「注目キーフレーム」という）が第 1 キーフレーム I 1、第 2 キーフレーム I 2、第 4 キーフレーム I 4、第 5 キーフレーム I 5 と特定される。第 1 キーフレーム I 1 と第 2 キーフレーム I 2 が第 1 画像対、第 4 キーフレーム I 4 と第 5 キーフレーム I 5 が第 2 画像対と定まる。つづいて、これら 4 つのキーフレームで形成される四辺形の中で中間フレーム I c が占める位置を幾何的に求める。つづいて、後述の補間によって中間フレームの画像を生成する。

## 【 0 1 7 9 】

このプロセスにおいて、中間フレームが図 2 0 において占める位置、および注目キーフレームは中間フレーム位置取得部 1 4 によって特定される。注目キーフレームはマッチングプロセッサ 1 8 に通知され、ここで第 1 画像対と第 2 画像対それぞれの間で前提技術にもとづくマッチング計算が行われる。それぞれのマッチング結果は対応点ファイルとして対応点ファイル保存部 2 0 へ記録される。

## 【 0 1 8 0 】

中間フレーム位置取得部 1 4 で取得された中間フレームの位置情報は中間フレーム生成部 2 2 に送られる。中間フレーム生成部 2 2 は、その位置情報とふたつの対応点ファイルをもとに補間計算を行う。

#### 【0 1 8 1】

図 2 1 は、補間の方法を示す。ここでは、第 1 キーフレーム I 1、第 2 キーフレーム I 2、第 4 キーフレーム I 4、第 5 キーフレーム I 5 をそれぞれ点 P 1、P 2、P 4、P 5 で模式的に示したとき、それらで定まる四辺形において、中間フレームを模式的に示す点 P c の位置が以下の条件を満たすとする。

#### 【0 1 8 2】

「P 1 と P 2 を結ぶ辺を  $s : (1 - s)$  に内分する点 Q と、P 4 と P 5 を結ぶ辺を  $s : (1 - s)$  に内分する点 R を結ぶ線分を  $(1 - t) : t$  に内分する」

中間フレーム生成部 2 2 は、第 1 画像対に関する対応点ファイルをもとにまず点 Q に当たる画像を  $s : (1 - s)$  の比による補間で生成する。つづいて、第 2 画像対に関する対応点ファイルをもとに点 R に当たる画像を  $s : (1 - s)$  の比による補間で生成する。最後に、これらふたつの画像を  $(1 - t) : t$  の比による補間で生成する。

中間フレーム生成部 2 2 は、第 1 画像対に関する対応点ファイルをもとにまず点 Q に当たる画像を  $s : (1 - s)$  の比による補間で生成する。つづいて、第 2 画像対に関する対応点ファイルをもとに点 R に当たる画像を  $s : (1 - s)$  の比による補間で生成する。最後に、これらふたつの画像を  $(1 - t) : t$  の比による補間で生成する。

図 2 2 は画像補間装置 1 0 の処理手順を示す。最初に表示部 2 4 には、任意のキーフレームが表示されている。ここでは仮に、図 2 0 の第 4 キーフレーム I 4 が表示されているとする。ユーザがマウスボタンを押しながら画面上で右上になぞると、表示されている商品（図示せず）が向かって右上の方向へ回転するよう指示されたと解釈する。したがって、図 2 0 のごとく、中間フレームの位置が第 4 キーフレーム I 4 から見て右上に移動する。マウスの移動方向と移動量をもとに、中間フレーム位置取得部 1 4 は図 2 0 または図 2 1 における中間フレームの位置を取得する（S 1 0 0 0）。

## 【0183】

つづいて、中間フレーム位置取得部14は注目キーフレームとして前述の4枚を選定し(S1002)、これをマッチングプロセッサ18へ伝える。マッチングプロセッサ18はそれらのキーフレームの画像をキーフレーム記憶部16から読み出し、それぞれ第1画像対と第2画像対についてマッチング計算を行う(S1004)。結果はふたつの対応点ファイルとして対応点ファイル保持部20へ格納される。

## 【0184】

中間フレーム生成部22は、それらの対応点ファイルからまず図21の点Q、Rを個別に求め、つづいて中間フレームを補間により求める(S1006)。最後に、ユーザのマウス操作の結果として、生成された中間フレームを表示する(S1008)。対応点ファイルは必要に応じてネットワークへ出力される。

## 【0185】

以上、実施の形態を説明した。ここでは四辺形をもとに補間を行ったが、これは三角形をもとに行ってもよい。また、対応点ファイルは都度生成したが、これは予めキーフレーム間について計算して生成しておいてもよい。当然その場合、中間フレームの生成がより高速に実現でき、リアルタイムの商品プレゼンテーションには最適である。

## 【図面の簡単な説明】

【図1】 図1(a)とは図1(b)は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図1(c)と図1(d)は、ふたりの人物の顔に関して前提技術で求められる $p(5, 0)$ の画像、図1(e)と図1(f)は、ふたりの人物の顔に関して前提技術で求められる $p(5, 1)$ の画像、図1(g)と図1(h)は、ふたりの人物の顔に関して前提技術で求められる $p(5, 2)$ の画像、図1(i)と図1(j)は、ふたりの人物の顔に関して前提技術で求められる $p(5, 3)$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図2】 図2(R)はもとの四辺形を示す図、図2(A)、図2(B)、図2(C)、図2(D)、図2(E)はそれぞれ相続四辺形を示す図である。

【図 3】 始点画像と終点画像の関係、および第  $m$  レベルと第  $m-1$  レベルの関係を相続四辺形を用いて示す図である。

【図 4】 パラメータ  $\eta$  とエネルギー  $C_f$  の関係を示す図である。

【図 5】 図 5 (a)、図 5 (b) は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図 6】 前提技術の全体手順を示すフローチャートである。

【図 7】 図 6 の S 1 の詳細を示すフローチャートである。

【図 8】 図 7 の S 1 0 の詳細を示すフローチャートである。

【図 9】 第  $m$  レベルの画像の一部と、第  $m-1$  レベルの画像の一部の対応関係を示す図である。

【図 1 0】 前提技術で生成された始点階層画像を示す図である。

【図 1 1】 図 6 の S 2 に進む前に、マッチング評価の準備の手順を示す図である。

【図 1 2】 図 6 の S 2 の詳細を示すフローチャートである。

【図 1 3】 第 0 レベルにおいて副写像を決定する様子を示す図である。

【図 1 4】 第 1 レベルにおいて副写像を決定する様子を示す図である。

【図 1 5】 図 1 2 の S 2 1 の詳細を示すフローチャートである。

【図 1 6】 ある  $f^{(m, s)}$  について  $\lambda$  を変えながら求められた  $f^{(m, s)}$  ( $\lambda = i \Delta \lambda$ ) に対応するエネルギー  $C^{(m, s)}_f$  の挙動を示す図である。

【図 1 7】  $\eta$  を変えながら求められた  $f^{(n)}$  ( $\eta = i \Delta \eta$ ) ( $i = 0, 1, \dots$ ) に対応するエネルギー  $C^{(n)}_f$  の挙動を示す図である。

【図 1 8】 図 1 8 (a) ~ 1 8 (e) は、実施の形態で生成すべき中間フレームを模式的に示す図である。

【図 1 9】 実施の形態に係る画像補間装置の構成図である。

【図 2 0】 実施の形態で生成すべき中間フレームとそのもとになるキーフレームの位置関係を概念的に示す図である。

【図 2 1】 実施の形態の画像補間装置で実施される補間処理の方法を示す図である。

【図 2 2】 実施の形態の画像補間装置の処理手順を示すフローチャートである。

【符号の説明】

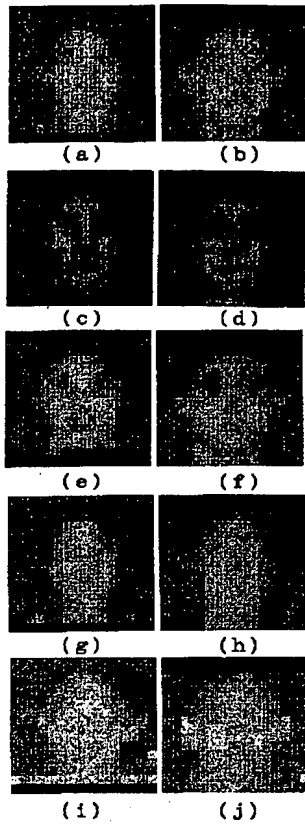
- 1 0 画像補間装置
- 1 2 G U I
- 1 4 中間フレーム位置取得部
- 1 6 キーフレーム記憶部
- 1 8 マッチングプロセッサ
- 2 0 対応点ファイル保持部
- 2 2 中間フレーム生成部
- 2 4 表示部
- 2 6 通信部



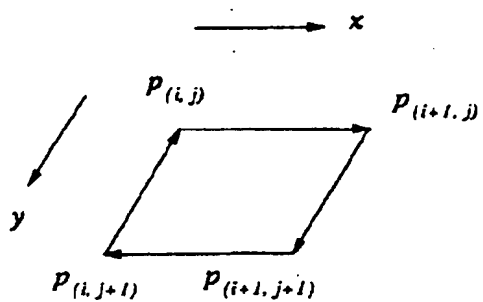
【書類名】

図面

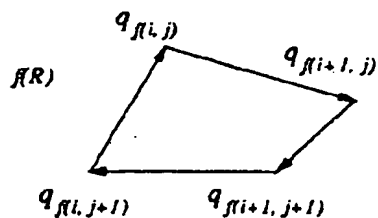
【図 1】



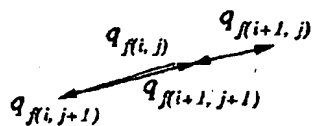
【図2】



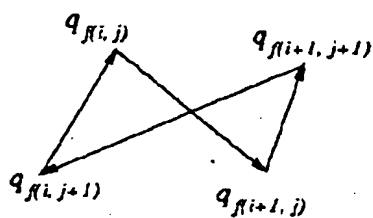
(R)



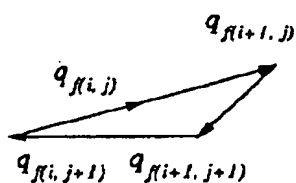
(A)



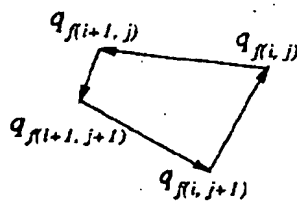
(E)



(B)

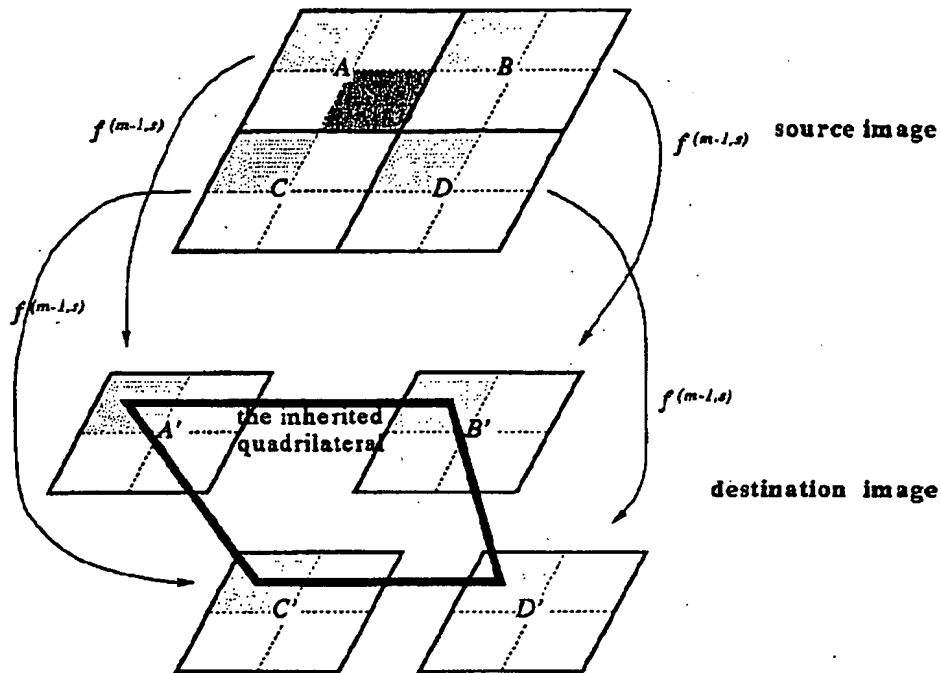


(D)

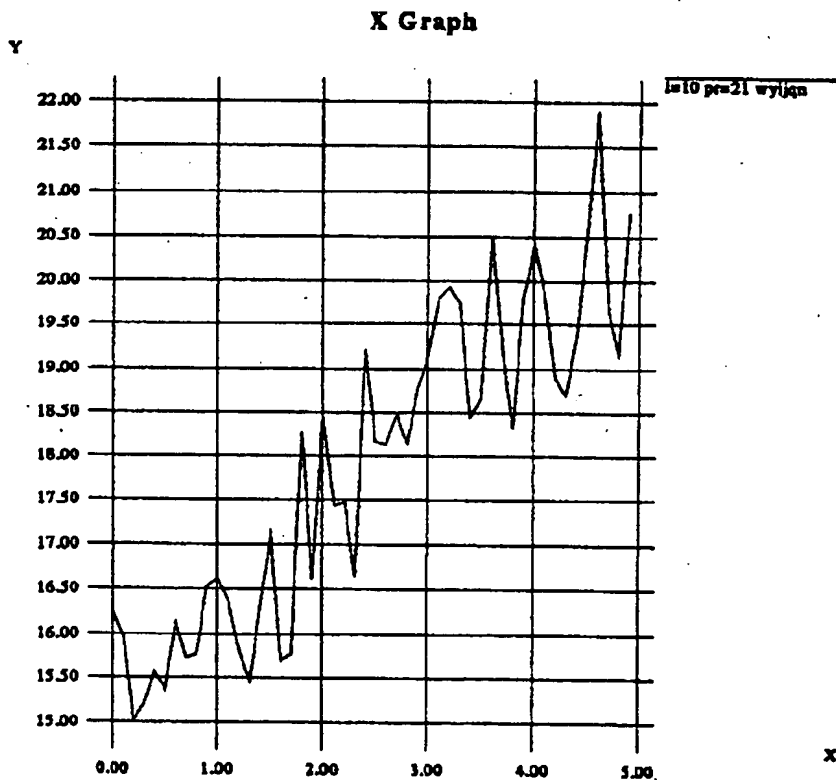


(C)

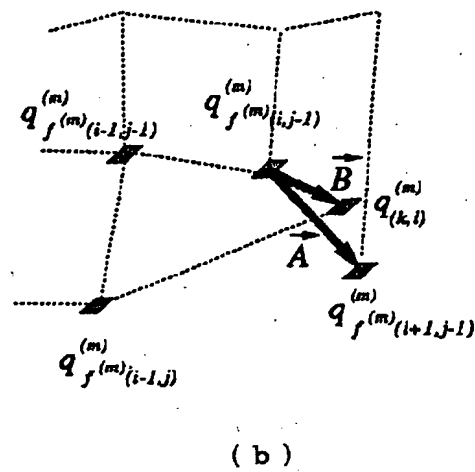
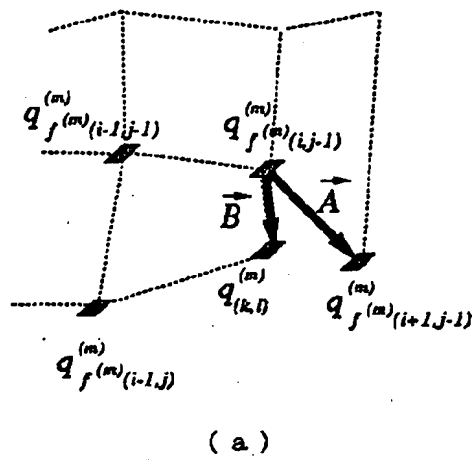
【図 3】



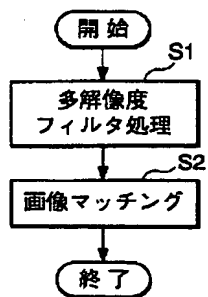
【図 4】



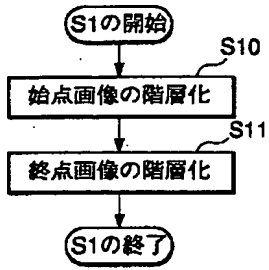
【図 5】



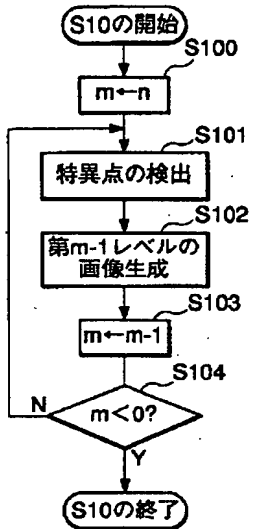
【図 6】



【図 7】



【図 8】



【図 9】

10	8		
6	3		

$p(m, a)$

3		

$p(m-1, 0)$

8		

$p(m-1, 1)$

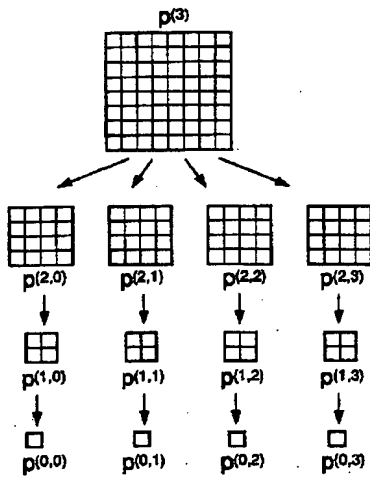
6		

$p(m-1, 2)$

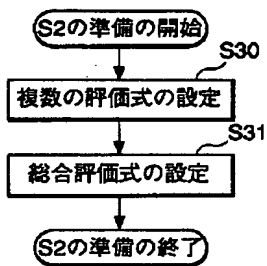
10		

$p(m-1, 3)$

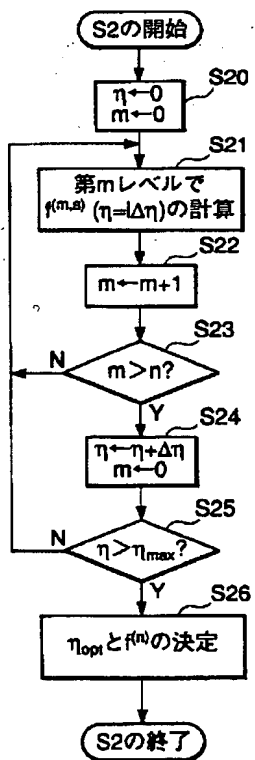
【図 10】



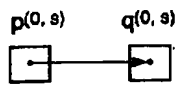
【図 11】



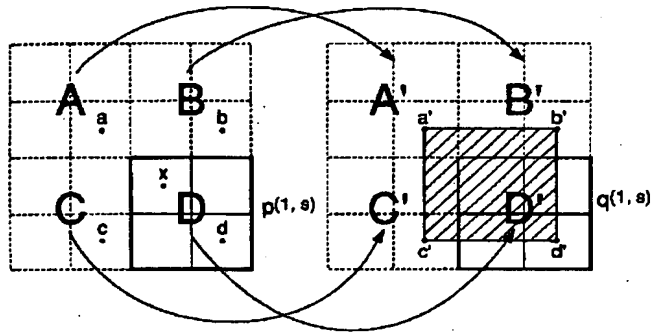
【図 12】



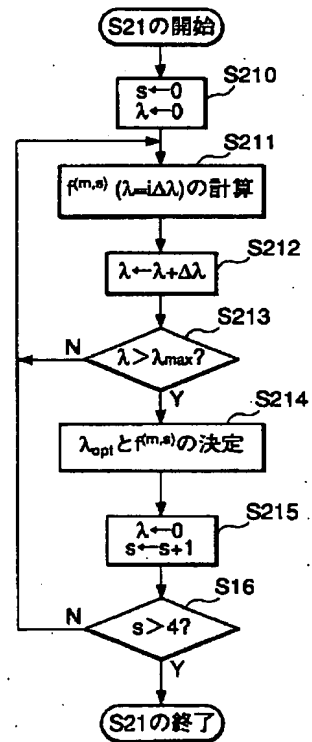
【図 13】



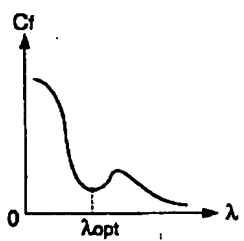
【図 14】



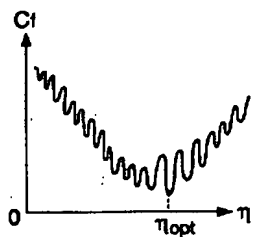
【図 15】



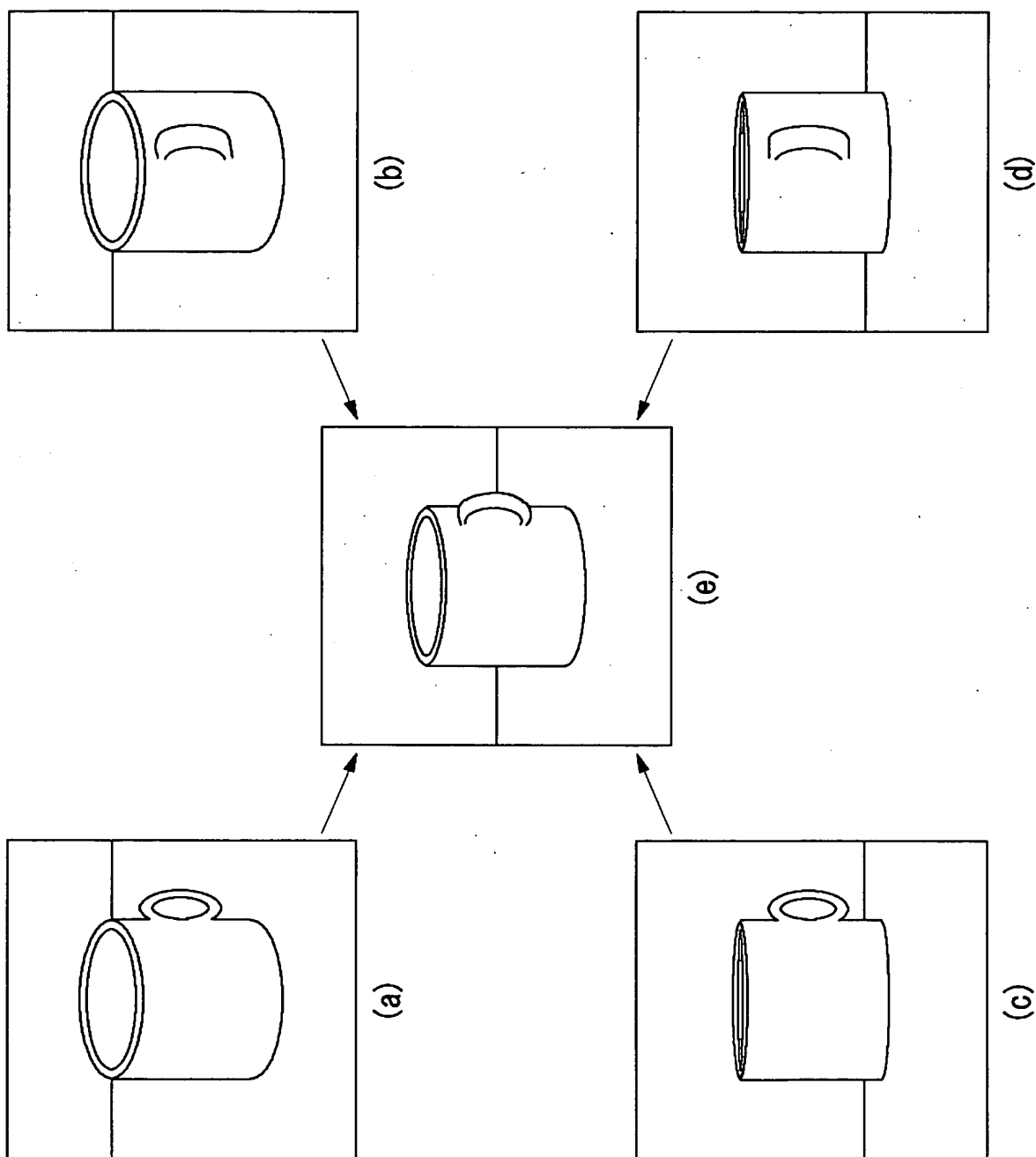
【図 16】



【図 17】

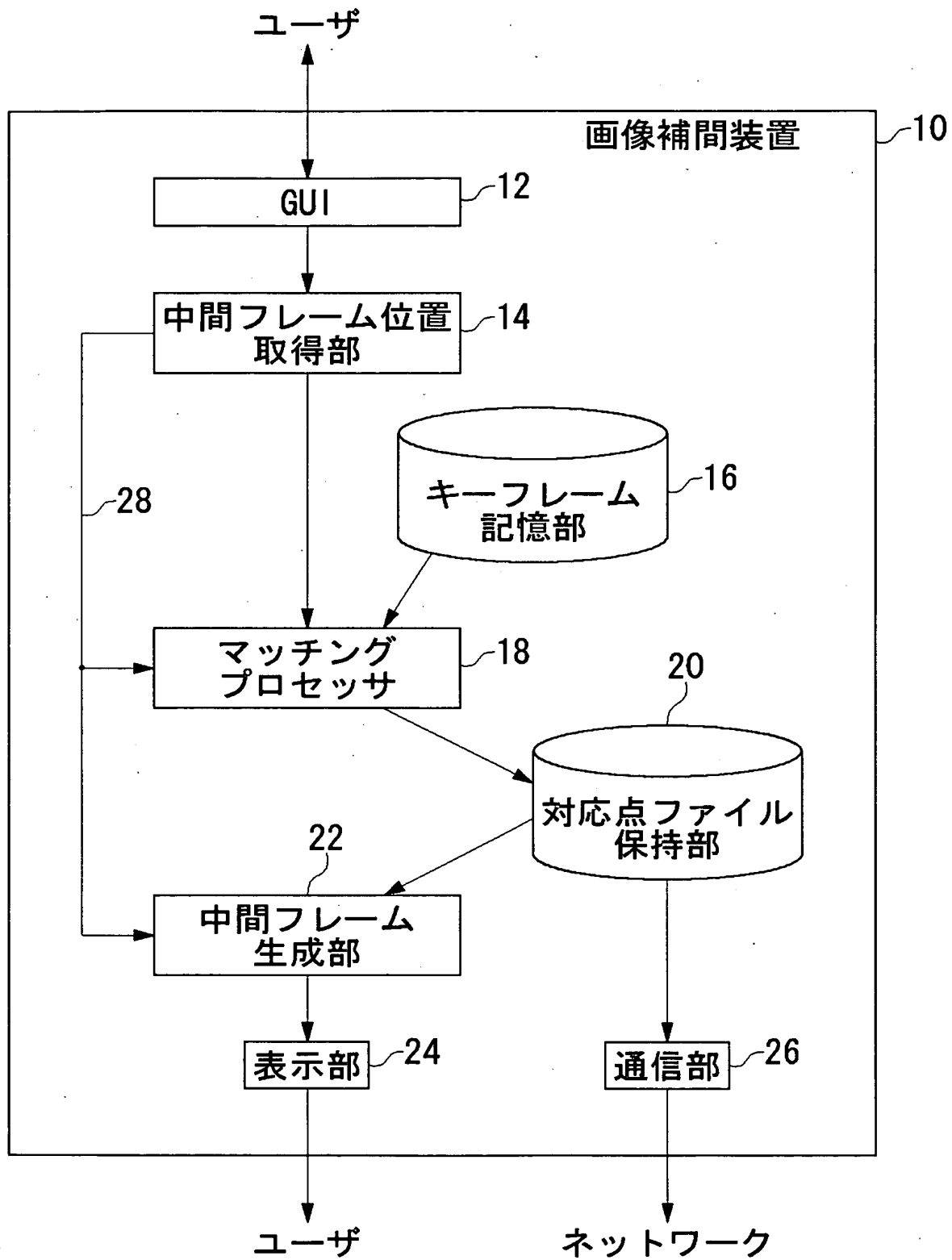


【図 18】

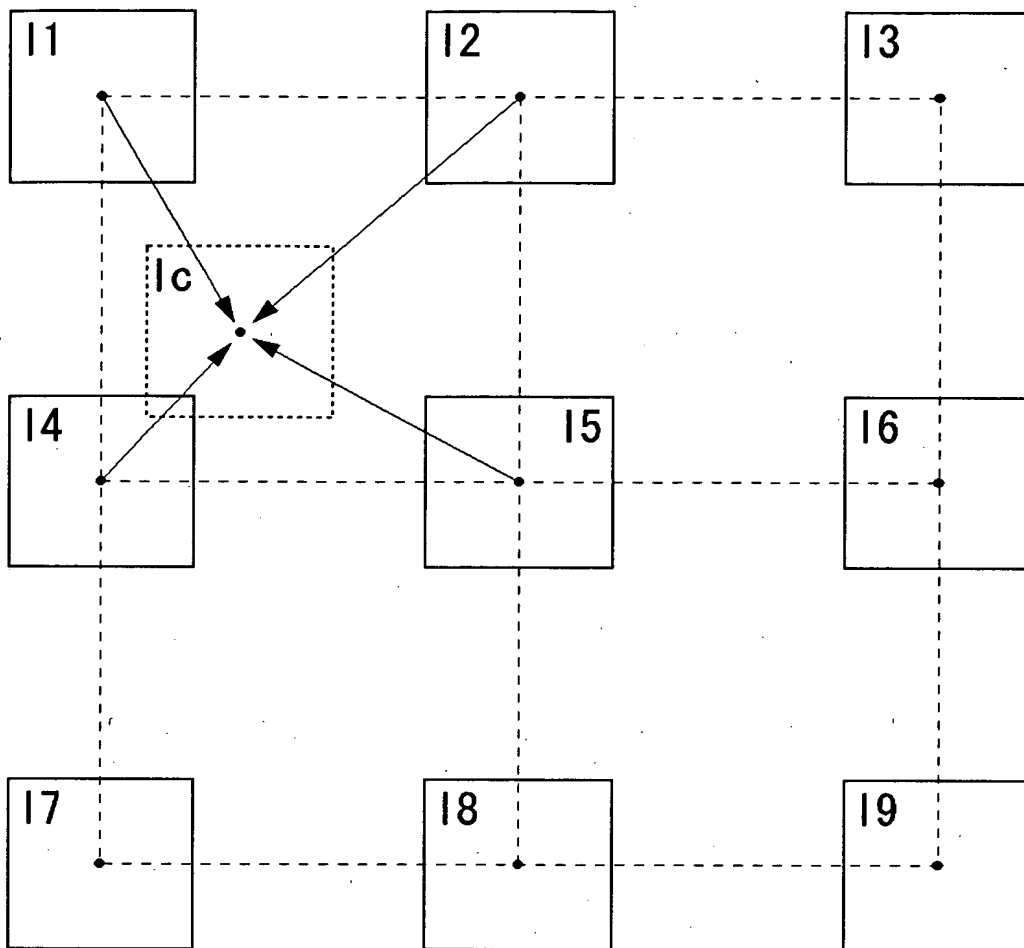




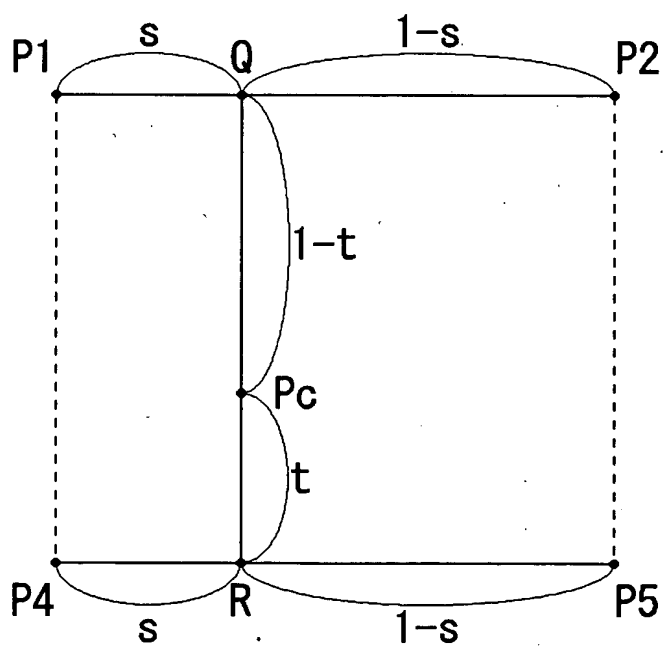
【図19】



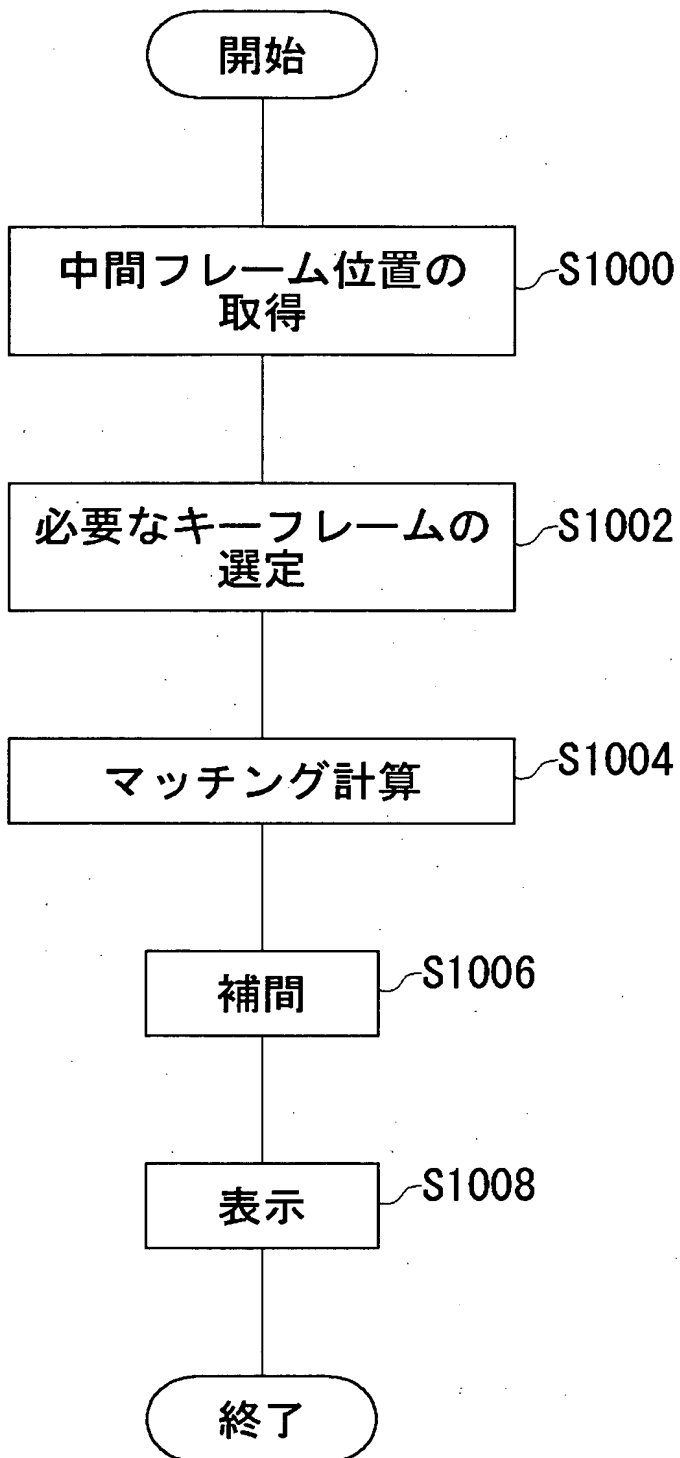
【図 20】



【図 21】



【図 2 2】



【書類名】 要約書

【要約】

【課題】 ECにおいて、商品の三次元プレゼンテーションを軽いデータ量でリアルに行いたい。

【解決手段】 キーフレーム記憶部 1 6 は、商品を縦方向および横方向に飛び飛びに撮影したキーフレームを保持する。中間フレーム位置取得部 1 4 は、生成すべき中間フレームの位置をキーフレームとの関係において取得する。マッチングプロセッサ 1 8 は、中間フレームを取り囲むキーフレーム間のマッチングを計算する。中間フレーム生成部 2 2 は、マッチング結果をもとに中間フレームを生成する。

【選択図】 図 1 9

出 願 人 履 歴 情 報

識別番号 [396001980]

1. 変更年月日 1996年 1月26日  
[変更理由] 新規登録  
住 所 東京都港区麻布十番1丁目7番3号  
氏 名 株式会社モノリス